

# Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US04/022605

International filing date: 14 July 2004 (14.07.2004)

Document type: Certified copy of priority document

Document details: Country/Office: US  
Number: 60/486,750  
Filing date: 14 July 2003 (14.07.2003)

Date of receipt at the International Bureau: 11 November 2004 (11.11.2004)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland  
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse



# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

*November 05, 2004*

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: 60/486,750

FILING DATE: *July 14, 2003*

RELATED PCT APPLICATION NUMBER: *PCT/US04/22605*

Certified by



Jon W Dudas

Acting Under Secretary of Commerce  
for Intellectual Property  
and Acting Director of the U.S.  
Patent and Trademark Office



Please type a plus sign (+) inside this box



Revised PTO/SB/16 (8-00)  
Approved for use through 10/31/2002. OMB 0651-0032  
Patent and Trademark Office, U.S. DEPARTMENT OF COMMERCE  
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.  
Attorney Docket No. 41860-189496

## PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53 (c).

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Evans	TARACHA	Nairobi, KENYA
Jane	GLEW	Nairobi, KENYA
Simon	GRAHAM	Nairobi, KENYA
Duncan	MWANGI	Nairobi, KENYA
Yoshikazu	HONDA	Nairobi, KENYA
Roger	PELLE	Nairobi, KENYA
John	TONUKARI	Nairobi, KENYA
Matasuke	YAMAGE	Nairobi, KENYA

☐ Additional inventors are being named on the \_\_\_\_\_ separately numbered sheets attached hereto

### TITLE OF THE INVENTION (280 characters max)

EAST COAST FEVER VACCINE BASED ON CTL-SPECIFIC SCHIZONT ANTIGENS

Direct all correspondence to:

### CORRESPONDENCE ADDRESS

☒ Customer Number

26694

OR

Type Customer Number here



26694

PATENT TRADEMARK OFFICE



Firm or  
Individual Name

VENABLE

Address

P.O. Box 34385

Address

City

Washington

State

DC

ZIP

20043-9998

Country

U.S.A.

Telephone

202.962.4800

Fax

202.962.8300

### ENCLOSED APPLICATION PARTS (check all that apply)



Specification Number of Pages

53



CD(s), Number



Drawing(s) Number of Sheets



Other (specify)



Application Data Sheet. See 37 CFR 1.76

Annex A, 68 pages, including Figs 1-23  
Annex B, 7 pages, including Figs 1-6

### METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT (check one)



Applicant claims small entity status. See 37 CFR 1.27.



A check or money order is enclosed to cover the filing fees

FILING FEE  
AMOUNT (\$)



The Commissioner is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number:

22-0261

80.00



Payment by credit card. Form PTO-2038 is attached.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are: \_\_\_\_\_

Respectfully submitted,  
SIGNATURE

Date

7/14/03

TYPED or PRINTED NAME Michael A. Gollin

REGISTRATION NO.  
(if appropriate)

31,957

Docket Number:

41860-189496

TELEPHONE (202) 962-4072

**VENABLE**  
ATTORNEYS AT LAW

17537 U.S. PTO

60/486750

07/14/03

# 1 EAST COAST FEVER VACCINE BASED ON CTL-SPECIFIC SCHIZONT ANTIGENS

2

## 3 BACKGROUND OF THE INVENTION

4 Theilerioses are a group of disease syndromes affecting cattle, sheep, goats and  
5 domestic buffalo caused by tick-borne haemo-protozoan parasites in the genus *Theileria*.  
6 The most economically important diseases include Mediterranean fever, East Coast fever  
7 (ECF) and Malignant theileriosis. Mediterranean fever caused by *Theileria annulata*  
8 occurs in North Africa, southern Europe, Near East, Middle East and many parts of Far  
9 East Asia with a population of 200 million cattle and buffalo at risk. ECF, caused by *T.*  
10 *parva*, affects 30 million cattle in eastern, central and southern Africa. Malignant  
11 theileriosis caused by *T. lestoquardi* affects sheep and goats in southeastern Europe,  
12 North Africa, the Near and Middle East and southern Russia and neighbouring States.  
13 These parasites belong to the same api-complex group as *Plasmodium falciparum*,  
14 *Toxoplasma gondii*, *Cytosporidium* spp, *Eimeria* spp and *Babesia* spp, with a life-cycle  
15 having the arthropod and mammalian components in which sexual and asexual stages  
16 develop, respectively. The pathogenic stages of *Theileria* parasites differ. *T. parva* causes  
17 a lympho-proliferative disorder in which schizont-infected lymphoblasts are responsible  
18 for the pathogenesis of the disease. On the other hand, anaemic disease caused by *T.*  
19 *lestoquardi* and *T. sergenti* is due to piroplasm-infected erythrocytes while both the  
20 schizont and piroplasm of *T. annulata* are pathogenic resulting in lympho-proliferative and  
21 anaemic syndromes, respectively.

22

23 Currently, theilerioses are controlled largely by tick control using acaricides and through  
24 "infection and treatment" vaccination protocols, of animals at risk. Due to cost and  
25 problems of tick resistance and environmental pollution, control of these diseases through  
26 acaricidal destruction of ticks is not sustainable. Vaccination, on the other hand, while  
27 effective presents with certain shortcomings associated with the use of live vaccines.  
28 Owing to the ease with which to transmit *T. annulata*, infected blood was originally used  
29 to immunise cattle with parasites of low virulence but were still accompanied by clinical  
30 episodes. With the advent of *in vitro* cultivation of *T. annulata* (Sharma et al., 1998) and  
31 the development of bulk culture techniques in the 1960s, significant progress was made  
32 in realising a practical immunisation strategy. Currently, passage-attenuated cultures of *T.*  
33 *annulata* are routinely used in national vaccination programs in affected countries. By  
34 contrast, similar efforts to immunise cattle against *T. parva* were unsuccessful. This was  
35 attributed to the failure of attenuated *T. parva* parasites to induce immunity. In addition,  
36 much higher numbers of *T. parva*-infected cells were required to infect cattle reliably since  
37 the schizonts of *T. parva* transfer at a low frequency and donor cells get rejected before  
38 successful transfer. *T. lestoquardi* has also been cultivated *in vitro* and studies have  
39 shown that attenuated parasites can be used to immunise animals with a degree of  
40 success.

41

42 Given the unsuccessful attempts to immunise cattle with attenuated *T. parva*, subsequent  
43 efforts have focussed on the use of virulent parasites with accompanying chemotherapy.  
44 The rationale of this infection and treatment method (ITM) is to allow the infection to  
45 establish and suppress development of patent clinical disease by administering  
46 theileriacidal drugs. Animals thus immunised were found to be protected against the



1 homologous parasite. This vaccination strategy has undergone successive refinement  
 2 including the use of cryopreserved triturated tick stabilites containing sporozoites (the  
 3 parasite stage infective for cattle lymphocytes) to standardise the infection dose, as well  
 4 as simultaneous drug administration. Further improvement of this immunisation approach  
 5 has involved the identification and use a combination of parasite stocks to broaden the  
 6 immunising spectrum of the vaccine against several field *T. parva* parasite populations.  
 7 But the use of local parasite stocks to immunise in areas where they have been isolated  
 8 is also practised. ITM immunisation against *T. parva* has been tested extensively under  
 9 laboratory and field conditions and is now deployed in the affected region to control ECF.

10  
 11 ITM is a very efficacious vaccine, however, it has a number of practical limitations that  
 12 hinder its application as a sustainable control measure against ECF. Being live, it requires  
 13 a cold chain, which is impractical in Africa, and also causes clinical disease if drug  
 14 application is inadequate and has the potential to introduce new parasite strains in areas  
 15 under the vaccination campaign. The cost (US\$10-20 per immunisation) this vaccine is  
 16 well beyond the poor farmers afflicted by ECF due to the cost of the drugs and the  
 17 requirement for a trained veterinarian to administer the vaccine. Because of these  
 18 concerns for the ITM vaccine, a great deal of investment has been put in research work to  
 19 develop a vaccine that will be sustainable.

20  
 21 Antigens of parasitic protozoans that induce a protective antibody response against the  
 22 development of disease, have been identified. For example, the major merozoite surface  
 23 protein of *Plasmodium* species has been shown to be a target of varying degrees of  
 24 protective immunity against the asexual blood stages in rodent and human malaria.  
 25 Vaccination of mice with purified P230, the major merozoite surface protein of the rodent  
 26 malaria *Plasmodium yoelii*, has resulted in reduced parasitemias in comparison to  
 27 controls upon intravenous challenge with a lethal dose of parasitized erythrocytes (Holder  
 28 et al. 1981. Nature 294:361). Mice have also been protected against *P. yoelii* by passive  
 29 transfer of a monoclonal antibody (Mab) specific for P230 (Majarian et al. 1984. J.  
 30 Immunol. 132:3131) and against (rodent malaria) *Plasmodium chabaudi adami* challenge  
 31 by passive immunization with a Mab specific for the homologous 250-kDa molecule of this  
 32 *Plasmodium* species (Lew et al. 1989. Proc. Natl. Acad. Sci. USA 86:3768).

33  
 34 A 67 kDa glycoprotein (p67) from the surface of the *T. parva* sporozoite has been isolated  
 35 (U.S. Patent Number 5273744). Cattle recovering from a single infection with *T. parva*  
 36 sporozoites resist homologous challenge. These animals have weak antibody and T cell  
 37 responses to p67. However, when repeatedly exposed to sporozoites, high anti-p67  
 38 antibodies are detected which neutralize infectivity of sporozoites. Cattle trials with this  
 39 candidate vaccine have demonstrated less than optimal protection despite induction of  
 40 strong antibody responses using traditional antigen delivery methods. Protection  
 41 engendered following a single infection has been shown to be based on the induction of  
 42 class I MHC-restricted CD8<sup>+</sup> cytotoxic T lymphocytes (CTLs) whose target parasite  
 43 antigens would be prime vaccine candidates. These studies, conducted over a period of  
 44 20 years, have made the development of an improved sub-unit vaccine against ECF a  
 45 reasonable probability. However, it has been impossible to identify the specific pathogen  
 46 antigens that could induce a CD8<sup>+</sup> cytotoxic T lymphocyte response. Methods using

1 traditional reagents and cell lines have failed to identify particular antigens that will  
2 stimulate this CD8+ lymphocyte pathway.

## 3 4 5 6 SUMMARY OF THE INVENTION

7  
8 The present invention relates to compositions and methods for the identification of  
9 parasite antigens, such as *Theileria parva* antigens, that trigger antigen-specific cytotoxic  
10 T lymphocyte (CTLs) responses, for inducing immunoprotection against *T. parva* in  
11 bovine species. More particularly, the method steps wherein the stimulation of  
12 responding lymphocytes to cells transfected by cDNAs encoding parasite antigen, is  
13 measured in a high throughput manner, by the release of soluble factors, such as gamma  
14 interferon, using either an antibody-elispot assay or a bioassay employing endothelial  
15 cells; the use of immortalized skin fibroblast cells from outbred animals, that have  
16 recovered from exposure, as antigen presenting cells, enables antigen identification,  
17 especially, where cloned bovine MHC class I genes are not available for co-transfection  
18 into COS cells; and/or the resolution of the identity of individual antigens from candidate  
19 cDNA pools.

## 20 21 BRIEF DESCRIPTION OF THE DRAWINGS

22  
23 FIG. 1 is a photograph on an ethidium bromide-stained agarose gel of total *T. parva*  
24 schizont RNA, poly(A+) RNA, and poly(A-)RNA.

25  
26 FIG. 2 is a photograph of an autoradiograph of 32P-labelled double stranded cDNA  
27 copied from poly(A+) *T. parva*, schizont RNA.

28  
29 FIG. 3 is a graph of Elispot data, indicating the mean spot forming cells/well when COS-7  
30 cells, co-transfected with a bovine MHC Class I gene and *T. parva* schizont cDNA pools,  
31 were co-cultured with cytotoxic T lymphocyte cell (CTL) lines, as measured by CTL  
32 gamma interferon secretion.

33  
34 FIG. 4 is a graph of Elispot data, indicating the mean spot forming cells/well when COS-7  
35 cells, co-transfected with a bovine MHC Class I gene and *T. parva* schizont cDNAs  
36 consisting of a pool of 10cDNAs, were co-cultured with cytotoxic T lymphocyte cell (CTL)  
37 lines, as measured by CTL gamma interferon secretion.

38  
39 FIG. 5 is a graph of Elispot data, indicating the mean spot forming cells/well when COS-7  
40 cells, co-transfected with a bovine MHC Class I gene and individual single *T. parva*  
41 schizont cDNA clones, were co-cultured with cytotoxic T lymphocyte cell (CTL) lines, as  
42 measured by CTL gamma interferon secretion.

43  
44 FIG. 6 is a graph of Elispot data, indicating the mean spot forming cells/well when COS-7  
45 cells, co-transfected with a bovine MHC Class I gene and an individual single *T. parva*  
46 schizont cDNA clone, Tp1, were co-cultured with cytotoxic T lymphocyte cell (CTL) lines,  
47 as measured by CTL gamma interferon secretion.

1  
2 FIG. 7 is a graph of Elispot data comparing the ability of COS-7 co-transfected cells to  
3 present target antigen to CTLs, to the ability of immortalized skin fibroblast cells (iSF) to  
4 present target antigen to CTLs.  
5  
6 FIG. 8(a) is a graph of Elispot data 9(b) is a graph of bioassay data. Both assays  
7 measure the secretion of gamma interferon by CTLs stimulated with Tp1.  
8  
9 FIG. 9 is a graph indicating the % lysis of COS-7 cells, co-transfected with a MHC Class I  
10 gene and a cDNA for Tp1.  
11  
12 FIG. 10 is the deduced amino acid sequence of Exonuclease III deleted Tp1 cDNA  
13 plasmids.  
14  
15 FIG. 11 is the deduced amino acid sequence of deleted Tp1 clones.  
16  
17 FIG. 12 is a graph of Elispot data indicating mapping of a Tp1 epitope using a synthetic  
18 peptide library.  
19  
20 FIG. 13 is a graph of Elispot data indicating mapping of the minimal length of the HD6  
21 restricted epitope of Tp1.  
22  
23 FIG. 14 is a comparison of the deduced amino acid sequences of Tp1 from two different  
24 T. parva strains, Muguga and Marikebuni.  
25  
26  
27  
28 FIG. 15 is a graph of Elispot data that indicates the response of CD8+ T cells, harvested  
29 from an immune bull to the HD6 restricted epitope of Tp1.  
30  
31 FIG. 16 is the DNA sequence and deduced protein sequence of Tp1.  
32  
33 FIG. 17 is the DNA sequence and deduced protein sequence of Tp4.  
34  
35 FIG. 18 is the DNA sequence and deduced protein sequence of Tp5.  
36  
37 FIG. 19 is the DNA sequence and deduced protein sequence of Tp7.  
38  
39 FIG. 20 is the DNA sequence and deduced protein sequence of Tp8.  
40  
41 FIG. 21(a) is a photograph of a Coomassie stained gel and an anti-His-tag immunoblot  
42 (Western blot) of Tp1 expression, using a bacterial expression vector; FIG. 22(b) is a  
43 photograph of a Coomassie stained gel and an anti-His-tag immunoblot of Tp4  
44 expression, using a bacterial expression vector; and, FIG. 21(b) is a photograph of a  
45 Coomassie stained gel and an anti-His-tag immunoblot of Tp5 expression, using a  
46 bacterial expression vector.

FIG. 22 is a panel of histograms indicating numbers of infected Bovine lymphocytes, stained with the reporter molecules, --mouse polyclonal antibody to Tp1 (FIG. 22(b)), mouse monoclonal antibody to Tp1 (FIG. 22(c)), mouse polyclonal antibody to Tp4 (FIG. 22(d)), and mouse polyclonal antibody to Tp5, (FIG. 22(e)).

FIG. 23. shows Tp1 multiple sequence alignments. Amino acid sequence comparison of a portion (containing the HD6 CTL epitope) of Tp1 generated from *T. parva* isolates from different regions. Domains with variations are underlined.

#### DETAILED DESCRIPTION OF THE INVENTION

Class I MHC-restricted CD8<sup>+</sup> cytotoxic T lymphocytes (CTLs) are responsible for protecting cattle against a lethal challenge with *T. parva* sporozoites (several reviews cited). These CTLs are directed at schizont-infected cells, which they recognize- and lyse. Schizont antigens that are recognized by these CD8<sup>+</sup> CTLs are the prime candidates for inclusion into an effective sub-unit vaccine for ECF.

The present invention provides polynucleotides, and methods for their identification, which encode useful proteins. The useful proteins encoded by the polynucleotide sequences of the subject invention are targets of antigen-specific cytotoxic CD8<sup>+</sup> T lymphocytes which have been shown to be protective in adoptive cell transfer experiments.

In a particularly preferred embodiment of the subject invention, the antigens encoded by polynucleotide sequences, cloned from expressed genes of *Theileria parva*, are identified by the use of immortalized cloned bovine skin fibroblast cell lines. In another preferred embodiment, stimulation of cytotoxic CD8<sup>+</sup> T lymphocytes, by antigen, is measured by the release of soluble factors, more specifically gamma interferon. The subject invention also relates to compositions isolated by methods including one or more of these particular steps and the use of these compositions: to stimulate or induce cytotoxic T cells, as diagnostic reagents for the detection of disease, or an immune response, in kits or high throughput "chip" methods for the detection of or expression of, identical or homologous nucleic acids.

As a result of the degeneracy of the genetic code, a multitude of nucleotide sequences may be produced which are based upon the sequences provided herein and corresponding peptides, polypeptides, or proteins. Some of these nucleotide sequences will bear only minimal homology to the sequences disclosed herein; however the subject invention specifically contemplates each and every possible variation of nucleotide

1 sequence that could be made by selecting combinations based on possible codon  
2 choices. These combinations are made in accordance with the standard triplet genetic  
3 code as applied to the nucleotide sequence of naturally occurring peptide, polypeptide, or  
4 protein, and all such variations are to be considered as being specifically disclosed  
5 herein.

6  
7 It is possible to produce the polynucleotides of the subject invention, or portions thereof,  
8 entirely by synthetic chemistry. After synthesis, the nucleic acid sequence can be used  
9 alone or joined with a preexisting sequence and inserted into one of the many available  
10 DNA vectors and their respective host cells using techniques well known in the art.  
11 Moreover, synthetic chemistry may be used to introduce specific mutations into the  
12 nucleotide sequence. Alternatively, a portion of sequence in which a mutation is desired  
13 can be synthesized and recombined with a portion of an existing genomic or recombinant  
14 sequence.

15  
16 Nucleotide sequences encoding a peptide, polypeptide, or protein may be joined to a  
17 variety of other nucleotide sequences by means of well established recombinant DNA  
18 techniques (Sambrook J. et al. (1989) Molecular Cloning: A Laboratory Manual, Cold  
19 Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; or Ausubel F. M. et al. (1989)  
20 Current Protocols in Molecular Biology, John Wiley & Sons, New York City). Useful  
21 sequences include an assortment of cloning vectors such as plasmids, cosmids, lambda  
22 phage derivatives, phagemids, and the like. Vectors of interest include vectors for  
23 replication, expression, probe generation, sequencing, and the like. In general, vectors of  
24 interest may contain an origin of replication functional in at least one organism,  
25 convenient restriction endonuclease sensitive sites, and selectable markers for one or  
26 more host cell systems.

27  
28 Another aspect of the subject invention is to provide for hybridization probes which are  
29 capable of hybridizing with naturally occurring antigen sequences or nucleotide  
30 sequences encoding the disclosed peptide, polypeptide, or protein. The stringency of the  
31 hybridization conditions will determine whether the probe identifies only the native  
32 nucleotide sequence or sequences of closely related molecules. If degenerate nucleotide  
33 sequences of the subject invention are used for the detection of related sequences, they  
34 should preferably contain at least 50% of the nucleotides of the sequences presented  
35 herein.

36  
37 Hybridization probes of the subject invention may be derived from the nucleotide  
38 sequences of the attached List Sequences and the Sequences provided in FIGS. 16-20,  
39 or from surrounding or included genomic sequences comprising untranslated regions  
40 such as promoters, enhancers and introns. Such hybridization probes may be labeled  
41 with appropriate reporter molecules. Means for producing specific hybridization probes  
42 include oligolabelling, nick translation, end-labeling or PCR amplification using a labeled  
43 nucleotide. Alternatively, the cDNA sequence may be cloned into a vector for the  
44 production of mRNA probe. Such vectors are known in the art, are commercially  
45 available, and may be used to synthesize RNA probes in vitro by addition of an  
46 appropriate RNA polymerase such as T7, T3 or SP6 and labelled nucleotides. A number



1 of companies (such as Pharmacia Biotech, Piscataway, N.J.; Promega, Madison, Wis.;  
2 US Biochemical Corp, Cleveland, Ohio; etc.) supply commercial kits and protocols for  
3 these procedures.

4  
5 The nucleotide sequences (shown in FIGS. 16-20,) can be used to generate probes for  
6 mapping the native genomic sequence. The sequence may be mapped to a particular  
7 chromosome or to a specific region of the chromosome using well known techniques.  
8 These include in situ hybridization to chromosomal spreads, flow-sorted chromosomal  
9 preparations, or artificial chromosome constructions such as yeast artificial chromosomes  
10 (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions or single  
11 chromosome cDNA libraries.

12  
13 In situ hybridization of chromosomal preparations and physical mapping techniques such  
14 as linkage analysis using established chromosomal markers are invaluable in extending  
15 genetic maps. The nucleotide sequences of the subject invention may also be used to  
16 detect differences in the chromosomal location of nucleotide sequences due to  
17 translocation, inversion, or recombination.

18  
19 Other aspects of the invention include use of the disclosed sequences or recombinant  
20 nucleic acids derived therefrom to produce purified peptides. The nucleotide sequences  
21 as disclosed herein may be used to produce an amino acid sequence using well known  
22 methods of recombinant DNA technology. Goeddel (Gene Expression Technology,  
23 Methods and Enzymology [1990] Vol 185, Academic Press, San Diego, Calif.) is one  
24 among many publications which teach expression of an isolated, purified nucleotide  
25 sequence. The amino acid or peptide may be expressed in a variety of host cells, either  
26 prokaryotic or eukaryotic. Host cells may be from the same species from which the  
27 nucleotide sequence was derived or from a different species.

28  
29 Still further aspects of the invention use these purified peptides to produce antibodies or  
30 other molecules able to bind to the peptides. These antibodies or binding agents can then  
31 be used for the screening of cells in order to localize the cellular distribution of the  
32 peptides or proteins. The antibodies are also useful for the affinity purification of  
33 recombinantly produced peptides or proteins. Such antibodies are also useful as  
34 diagnostic reagents for the detection of protozoan diseases or in antibody-mediated tests  
35 and assays.

36  
37 The disclosed nucleotide sequences can be used individually, or in panels, in tests or  
38 assays to detect levels of peptide, polypeptide, or protein expression. The form of such  
39 qualitative or quantitative methods may include northern analysis, dot blot or other  
40 membrane based technologies, dip stick, pin or chip technologies, PCR, ELISAs or other  
41 multiple sample format technologies.

42  
43 As used herein, the following definitions apply:

44  
45 An "oligonucleotide" or "oligomer" is a stretch of nucleotide residues which has a  
46 sufficient number of bases to be used in a polymerase chain reaction (PCR). These short

1 sequences are based on (or designed from) genomic or cDNA sequences and are used  
2 to amplify, confirm, or reveal the presence of an identical, similar or complementary DNA  
3 or RNA in a particular cell or tissue. Oligonucleotides or oligomers comprise portions of a  
4 DNA sequence having at least about 10 nucleotides and as many as about 50  
5 nucleotides, preferably about 15 to 30 nucleotides. They can be chemically synthesized  
6 and may be used as probes.

7

8 By the term "recombinant nucleic acid" herein is meant nucleic acid, originally formed in  
9 vitro or in a cell in culture, in general, by the manipulation of nucleic acid by  
10 endonucleases and/or exonucleases and/or polymerases and/or ligases and/or  
11 recombinases, to produce a nucleic acid not normally found in nature. Thus an isolated  
12 nucleic acid, in a linear form, or an expression vector formed in vitro by ligating DNA  
13 molecules that are not normally joined, are both considered recombinant for the purposes  
14 of this invention. It is understood that once a recombinant nucleic acid is made and  
15 reintroduced into a host cell or organism, it will replicate non-recombinantly, i.e., using the  
16 in vivo cellular machinery of the host cell rather than in vitro manipulations; however, such  
17 nucleic acids, once produced recombinantly, although subsequently replicated non-  
18 recombinantly, are still considered recombinant for the purposes of the invention.

19

20 "Probes" are nucleic acid sequences of variable length, preferably between at least about  
21 10 and as many as about 6,000 nucleotides, depending on use. They are used in the  
22 detection of identical, similar, or complementary nucleic acid sequences. Longer length  
23 probes are usually obtained from a natural or recombinant source, are highly specific and  
24 much slower to hybridize than oligomers. They may be single- or double-stranded and  
25 designed to have specificity in PCR, hybridization membrane-based, or ELISA-like  
26 technologies.

27

28 "Reporter" molecules are chemical moieties used for labeling a nucleic or amino acid  
29 sequence. They include, but are not limited to, radionuclides, enzymes, fluorescent,  
30 chemi-luminescent, or chromogenic agents. Reporter molecules associate with, establish  
31 the presence of, and may allow quantification of a particular nucleic or amino acid  
32 sequence.

33

34 A "portion" or "fragment" of a polynucleotide or nucleic acid comprises all or any part of  
35 the nucleotide sequence having fewer nucleotides than about 6 kb, preferably fewer than  
36 about 1 kb which can be used as a probe. Such probes may be labeled with reporter  
37 molecules using nick translation, Klenow fill-in reaction, PCR or other methods well  
38 known in the art. After pretesting to optimize reaction conditions and to eliminate false  
39 positives, nucleic acid probes may be used in Southern, northern or in situ hybridizations  
40 to determine whether target DNA or RNA is present in a biological sample, cell type,  
41 tissue, organ or organism.

42

43 A "polypeptide" comprises a protein, oligopeptide or peptide fragments thereof.

44

1 A mutant, variant, or modified polypeptide means any polypeptide encoded by a  
2 nucleotide sequence that has been mutated through insertions, deletions, substitutions, or  
3 the like.

4  
5 "Recombinant nucleotide variants" are alternate polynucleotides which encode a  
6 particular protein. They may be synthesized, for example, by making use of the  
7 "redundancy" in the genetic code. Various codon substitutions, such as the silent changes  
8 which produce specific restriction sites or codon usage-specific mutations, may be  
9 introduced to optimize cloning into a plasmid or viral vector or expression in a particular  
10 prokaryotic or eukaryotic host system, respectively.

11  
12 "Linkers" are synthesized palindromic nucleotide sequences which create internal  
13 restriction endonuclease sites for ease of cloning the genetic material of choice into  
14 various vectors. "Polylinkers" are engineered to include multiple restriction enzyme sites  
15 and provide for the use of both those enzymes which leave 5' and 3' overhangs such as  
16 BamHI, EcoRI, PstI, KpnI and Hind III or which provide a blunt end such as EcoRV,  
17 SnaBI and Stul.

18  
19  
20 "Chimeric" molecules are polynucleotides or polypeptides which are created by combining  
21 one or more nucleotide peptide sequences (or their parts). In the case of nucleotide  
22 sequences, such combined sequences may be introduced into an appropriate vector and  
23 expressed to give rise to a chimeric polypeptide which may be expected to be different  
24 from the native molecule in one or more of the following characteristics: cellular location,  
25 distribution, ligand-binding affinities, interchain affinities, degradation/turnover rate,  
26 signaling, etc.

27  
28 "Active" is that state which is capable of being useful or of carrying out some role. It  
29 specifically refers to those forms, fragments, or domains of an amino acid sequence  
30 which display the biologic and/or immunogenic activity characteristic of the naturally  
31 occurring peptide, polypeptide, or protein.

32  
33 "Naturally occurring" refers to a polypeptide produced by cells which have not been  
34 genetically engineered or which have been genetically engineered to produce the same  
35 sequence as that naturally produced.

36  
37 "Derivative" refers to those polypeptides which have been chemically modified by such  
38 techniques as ubiquitination, labeling, pegylation (derivatization with polyethylene glycol),  
39 and chemical insertion or substitution of amino acids such as ornithine which do not  
40 normally occur in proteins.

41  
42 "Recombinant polypeptide variant" refers to any polypeptide which differs from naturally  
43 occurring peptide, polypeptide, or protein by amino acid insertions, deletions and/or  
44 substitutions.

45  
46 Amino acid "substitutions" are defined as one for one amino acid replacements. They are

1 conservative in nature when the substituted amino acid has similar structural and/or  
2 chemical properties. Examples of conservative replacements are substitution of a leucine  
3 with an isoleucine or valine, an aspartate with a glutamate, or a threonine with a serine.

4  
5 Amino acid "insertions" or "deletions" are changes to or within an amino acid sequence.  
6 They typically fall in the range of about 1 to 5 amino acids. The variation allowed in a  
7 particular amino acid sequence may be experimentally determined by producing the  
8 peptide synthetically or by systematically making insertions, deletions, or substitutions of  
9 nucleotides in the sequence using recombinant DNA techniques.

10  
11  
12 An "oligopeptide" is a short stretch of amino acid residues and may be expressed from an  
13 oligonucleotide. Such sequences comprise a stretch of amino acid residues of at least  
14 about 5 amino acids and often about 17 or more amino acids, typically at least about 9 to  
15 13 amino acids, and of sufficient length to display biologic and/or immunogenic activity.

16  
17  
18 [0061] A "standard" is a quantitative or qualitative measurement for comparison.  
19 Preferably, it is based on a statistically appropriate number of samples and is created to  
20 use as a basis of comparison when performing diagnostic assays, running clinical trials,  
21 or following patient treatment profiles. The samples of a particular standard may be  
22 normal or similarly abnormal.

23  
24 An antibody or "specific binding parts" means any fragment of an antibody molecule that  
25 will bind antigen or other ligands such as lectins or other molecules. "Specific binding  
26 parts" is meant to include, but not be limited to antibody fragments such as Fab  
27 fragments, Fab'(2) fragments, Fc region fragments, Complementarity determining regions  
28 (CDRs), Fv fragments, single chain Fv (scFv) fragments, and antigen binding site  
29 fragments.

30  
31 A T cell "epitope" is an antigenic determinant recognized and bound by the T-cell  
32 receptor. Epitopes recognized by the T-cell receptor are often located in the inner,  
33 unexposed side of the antigen, and become accessible to the T-cell receptors after  
34 proteolytic processing of the antigen.

35  
36 An "ELISpot Assay: Short for Enzyme-linked ImmunoSpot Assay", means an antibody  
37 based method to detect secretion of soluble factors released by cells. Originally  
38 developed as a method to detect antibody-secreting B-cells, later the method was  
39 adapted to determine T-cell reaction to a specific antigen, often represented as number  
40 of activated cells per million.

41  
42 A Bovine MHC class II bioassay, means an assay measures IFN- $\gamma$  release from T cells  
43 responding to specific stimulation through the ability of IFN- $\gamma$  to induce and up-regulate  
44 expression of class II molecules on bovine endothelial cells. Bovine endothelial cells do  
45 not constitutively express class II molecules unless triggered by external signals such as  
46 IFN- $\gamma$ .

1  
2 Since the list of technical and scientific terms cannot be all encompassing, any undefined  
3 terms shall be construed to have the same meaning as is commonly understood by one  
4 of skill in the art to which this invention belongs. Furthermore, the singular forms "a", "an"  
5 and "the" include plural referents unless the context clearly dictates otherwise.

6  
7 The invention is not to be limited only to the particular sequences, variants, formulations  
8 or methods described. The sequences, variants, formulations and methodologies may  
9 vary, and the terminology used herein is for the purpose of describing particular  
10 embodiments. The terminology and definitions are not intended to be limiting.

11  
12 Polynucleotide probes employing the present inventive compositions. DNA possesses a  
13 fundamental property called base complementarity. In nature, DNA ordinarily exists in the  
14 form of pairs of anti-parallel strands, the bases on each strand projecting from that strand  
15 toward the opposite strand. The base adenine (A) on one stand will always be opposed  
16 to the base thymine (T) on the other strand, and the base guanine (G) will be opposed to  
17 the base cytosine (C). The bases are held in apposition by their ability to hydrogen bond  
18 in this specific way. Though each individual bond is relatively weak, the net effect of  
19 many adjacent hydrogen bonded bases, together with base stacking effects, is a stable  
20 joining of the two complementary strands. These bonds can be broken by treatments  
21 such as high pH or high temperature, and these conditions result in the dissociation, or  
22 "denaturation," of the two strands. If the DNA is then placed in conditions which make  
23 hydrogen bonding of the bases thermodynamically favorable, the DNA strands will  
24 anneal, or "hybridize," and reform the original double-stranded DNA. If carried out under  
25 appropriate conditions, this hybridization can be highly specific. That is, only strands with  
26 a high degree of base complementarity will be able to form stable double-stranded  
27 structures. The relationship of the specificity of hybridization to reaction conditions is well  
28 known. Thus, hybridization may be used to test whether two pieces of DNA are  
29 complementary in their base sequences. It is this hybridization mechanism which  
30 facilitates the use of probes of the subject invention to readily detect and characterize  
31 DNA sequences of interest.

32  
33 The specifically exemplified polynucleotides of the subject invention can themselves be  
34 used as probes. Additional polynucleotide sequences can be added to the ends of (or  
35 internally in) the exemplified polynucleotide sequences so that polynucleotides that are  
36 longer than the exemplified polynucleotides can also be used as probes. Thus, isolated  
37 polynucleotides comprising one or more of the exemplified sequences are within the  
38 scope of the subject invention. Polynucleotides that have less nucleotides than the  
39 exemplified polynucleotides can also be used and are contemplated within the scope of  
40 the present invention. For example, for some purposes, it might be useful to use a  
41 conserved sequence from an exemplified polynucleotide wherein the conserved  
42 sequence comprises a portion of an exemplified sequence. Thus, polynucleotides of the  
43 subject invention can be used to find additional, homologous (wholly or partially) genes.

44  
45 Probes of the subject invention may be composed of DNA, RNA, or PNA (peptide nucleic  
46 acid). The probe will normally have at least about 10 bases, more usually at least about



1 17 bases, and may have up to about 100 bases or more. Longer probes can readily be  
2 utilized, and such probes can be, for example, several kilobases in length. The probe  
3 sequence is designed to be at least substantially complementary to a portion of a gene  
4 encoding a protein of interest. The probe need not have perfect complementarity to the  
5 sequence to which it hybridizes. The probes may be labeled utilizing techniques that are  
6 well known to those skilled in this art.

7  
8 One approach for the use of the subject invention as probes entails first identifying DNA  
9 segments that are homologous with the disclosed nucleotide sequences using, for  
10 example, Southern blot analysis of a gene bank. Thus, it is possible, without the aid of  
11 biological analysis, to know in advance the probable activity of many new  
12 polynucleotides, and of the individual gene products expressed by a given  
13 polynucleotide. Such an analysis provides a rapid method for identifying commercially  
14 valuable compositions.

15  
16 One hybridization procedure useful according to the subject invention typically includes  
17 the initial steps of isolating the DNA sample of interest and purifying it chemically. Either  
18 lysed cells or total fractionated nucleic acid isolated from cells can be used. Cells can be  
19 treated using known techniques to liberate their DNA (and/or RNA). The DNA sample  
20 can be cut into pieces with an appropriate restriction enzyme. The pieces can be interest  
21 can be through electrophoresis in a gel, usually agarose or acrylamide. The pieces of  
22 interest can be transferred to an immobilizing membrane.

23  
24 The particular hybridization technique is not essential to the subject invention. As  
25 improvements are made in hybridization techniques, they can be readily applied.

26  
27 The probe and sample can then be combined in a hybridization buffer solution and held  
28 at an appropriate temperature until annealing occurs. Thereafter, the membrane is  
29 washed free of extraneous materials, leaving the sample and bound probe molecules  
30 typically detected and quantified by autoradiography and/or liquid scintillation counting.  
31 As is well known in the art, if the probe molecule and nucleic acid sample hybridize by  
32 forming a strong non-covalent bond between the two molecules, it can be reasonably  
33 assumed that the probe and sample are essentially identical or very similar. The probe's  
34 detectable label provides a means for determining in a known manner whether  
35 hybridization has occurred.

36  
37 In the use of the nucleotide segments as probes, the particular probe is labeled with any  
38 suitable label known to those skilled in the art, including radioactive and non-radioactive  
39 labels. Typical radioactive labels include <sup>32</sup>P, <sup>35</sup>S, or the like. Non-radioactive  
40 labels include, for example, ligands such as biotin or thyroxine, as well as enzymes such  
41 as hydrolases or peroxidases, or the various chemiluminescers such as luciferin, or  
42 fluorescent compounds like fluorescein and its derivatives. In addition, the probes can be  
43 made inherently fluorescent as described in International Application No. WO 93/16094.

44  
45 Various degrees of stringency of hybridization can be employed. The more stringent the  
46 conditions, the greater the complementarity that is required for duplex formation.

Stringency can be controlled by temperature, probe concentration, probe length, ionic strength, time, and the like. Preferably, hybridization is conducted under moderate to high stringency conditions by techniques well known in the art, as described, for example, in Keller, G. H., M. M. Manak (1987) DNA Probes, Stockton Press, New York, N.Y., pp. 169-170.

As used herein "moderate to high stringency" conditions for hybridization refers to conditions that achieve the same, or about the same, degree of specificity of hybridization as the conditions "as described herein." Examples of moderate to high stringency conditions are provided herein. Specifically, hybridization of immobilized DNA on Southern blots with <sup>32</sup>P-labeled gene-specific probes was performed using standard methods (Maniatis et al.). In general, hybridization and subsequent washes were carried out under moderate to high stringency conditions that allowed for detection of target sequences with homology to sequences exemplified herein. For double-stranded DNA gene probes, hybridization was carried out overnight at 20-25.degree. C. below the melting temperature (T<sub>m</sub>) of the DNA hybrid in 6.times. SSPE, 5.times. Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. The melting temperature is described by the following formula from Beltz et al. (1983):

$$T_m = 81.5 \text{ degree. C.} + 16.6 \log[\text{Na}^+] + 0.41(\% \text{ G+C}) - 0.61(\% \text{ formamide}) - 600/\text{length of duplex in base pairs.}$$

Washes are typically carried out as follows:

(1) Twice at room temperature for 15 minutes in 1.times. SSPE, 0.1% SDS (low stringency wash).

(2) Once at T<sub>m</sub>-20.degree. C. for 15 minutes in 0.2.times. SSPE, 0.1% SDS (moderate stringency wash).

For oligonucleotide probes, hybridization was carried out overnight at 10-20.degree. C. below the melting temperature (T<sub>m</sub>) of the hybrid in 6.times. SSPE, 5.times. Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. T<sub>m</sub> for oligonucleotide probes was determined by the following formula from Suggs et al. (1981):

$$T_m \text{ (.degree. C.)} = 2(\text{number T/A base pairs}) + 4(\text{number G/C base pairs})$$

Washes were typically carried out as follows:

(1) Twice at room temperature for 15 minutes 1.times. SSPE, 0.1% SDS (low stringency wash).

(2) Once at the hybridization temperature for 15 minutes in 1.times. SSPE, 0.1% SDS (moderate stringency wash).

In general, salt and/or temperature can be altered to change stringency. With a labeled

1 DNA fragment of greater than about 70 or so bases in length, the following conditions  
2 can be used:

3  
4 1 Low: 1 or 2X SSPE, room temperature Low: 1 or 2X SSPE, 42.degree. C. Moderate:  
5 0.2X or 1X SSPE, 65.degree. C. High: 0.1X SSPE, 65.degree. C.

6  
7 Duplex formation and stability depend on substantial complementarity between the two  
8 strands of a hybrid, and, as noted above, a certain degree of mismatch can be tolerated.  
9 Therefore, polynucleotide sequences of the subject invention include mutations (both  
10 single and multiple), deletions, and insertions in the described sequences, and  
11 combinations thereof, wherein said mutations, insertions, and deletions permit formation  
12 of stable hybrids with a target polynucleotide of interest. Mutations, insertions, and  
13 deletions can be produced in a given polynucleotide sequence using standard methods  
14 known in the art. Other methods may become known in the future.

15  
16 The mutational, insertional, and deletional variants of the polypeptide sequences of the  
17 invention can be used in the same manner as the exemplified polynucleotide sequences  
18 so long as the variants have substantial sequence similarity with the original sequence.  
19 As used herein, substantial sequence similarity refers to the extent of nucleotide  
20 similarity that is sufficient to enable the variant polynucleotide to function in the same  
21 capacity as the original sequence. Preferably, this similarity is greater than 50%; more  
22 preferably, this similarity is greater than 75%; and most preferably, this similarity is  
23 greater than 90%. The degree of similarity needed for the variant to function in its  
24 intended capacity will depend upon the intended use of the sequence. It is well within the  
25 skill of a person trained in this art to make mutational, insertional, and deletional  
26 mutations that are designed to improve the function of the sequence or otherwise provide  
27 a methodological advantage.

28  
29 PCR technology. Polymerase Chain Reaction (PCR) is a repetitive, enzymatic, primed  
30 synthesis of a nucleic acid sequence. This procedure is well known and commonly used  
31 by those skilled in this art (see U.S. Pat. Nos. 4,683,195, 4,683,202, and 4,800,159; Saiki  
32 et al., 1985). PCR is based on the enzymatic amplification of a DNA fragment of interest  
33 that is flanked by two oligonucleotide primers that hybridize to opposite strands of the  
34 target sequence. The primers are oriented with the 3' ends pointing towards each other.  
35 Repeated cycles of heat denaturation of the template, annealing of the primers to their  
36 complementary sequences, and extension of the annealed primers with a DNA  
37 polymerase result in the amplification of the segment defined by the 5' ends of the PCR  
38 primers. Since the extension product of each primer can serve as a template for the other  
39 primer, each cycle essentially doubles the amount of DNA fragment produced in the  
40 previous cycle. This results in the exponential accumulation of the specific target  
41 fragment, up to several million-fold in a few hours. By using a thermostable DNA  
42 polymerase such as Taq polymerase, which is isolated from the thermophilic bacterium  
43 *Thermus aquaticus*, the amplification process can be completely automated. Other  
44 enzymes that can be used are known to those skilled in the art.

45  
46 The polynucleotide sequences of the subject invention (and portions thereof such as

1 conserved regions and portions that serve to distinguish these sequences from  
2 previously-known sequences) can be used as, and/or used in the design of, primers for  
3 PCR amplification. In performing PCR amplification, a certain degree of mismatch can be  
4 tolerated between primer and template. Therefore, mutations, deletions, and insertions  
5 (especially additions of nucleotides to the 5' end) of the exemplified polynucleotides can  
6 be used in this manner. Mutations, insertions and deletions can be produced in a given  
7 primer by methods known to an ordinarily skilled artisan.

8  
9 Full length genes may be cloned utilizing partial nucleotide sequence and various  
10 methods known in the art. Gobinda et al. (1993; PCR Methods Applic 2:318-22) disclose  
11 "restriction-site PCR" as a direct method which uses universal primers to retrieve  
12 unknown sequence adjacent to a known locus. First, genomic DNA is amplified in the  
13 presence of primer to linker and a primer specific to the known region. The amplified  
14 sequences are subjected to a second round of PCR with the same linker primer and  
15 another specific primer internal to the first one. Products of each round of PCR are  
16 transcribed with an appropriate RNA polymerase and sequenced using reverse  
17 transcriptase.

18  
19 Inverse PCR can be used to acquire unknown sequences starting with primers based on  
20 a known region (Triglia T. et al. (1988) Nucleic Acids Res 16:8186). The method uses  
21 several restriction enzymes to generate a suitable fragment in the known region of a  
22 gene. The fragment is then circularized by intramolecular ligation and used as a PCR  
23 template. Divergent primers are designed from the known region. The multiple rounds of  
24 restriction enzyme digestions and ligations that are necessary prior to PCR make the  
25 procedure slow and expensive (Gobinda et al. [1993] supra).

26  
27 Capture PCR (Lagerstrom M. et al. (1991) PCR Methods Applic 1:111-19) is a method  
28 for PCR amplification of DNA fragments adjacent to a known sequence in eucaryotic and  
29 YAC DNA. As noted by Gobinda et al. (1993,supra), capture PCR also requires multiple  
30 restriction enzyme digestions and ligations to place an engineered double-stranded  
31 sequence into an unknown portion of the DNA molecule before PCR. Although the  
32 restriction and ligation reactions are carried out simultaneously, the requirements for  
33 extension, immobilization and two rounds of PCR and purification prior to sequencing  
34 render the method cumbersome and time consuming.

35  
36 Parker J. D. et al. (Nucleic Acids Res [1991 119:3055-60), teach walking PCR, a method  
37 for targeted gene walking which permits retrieval of unknown sequences.  
38 PromoterFinder.TM. is a kit available from Clontech Laboratories, Inc. (Palo Alto, Calif.)  
39 which uses PCR and primers derived from p53 to walk in genomic DNA. Nested primers  
40 and special PromoterFinder.TM. libraries are used to detect upstream sequences such  
41 as promoters and regulatory elements. This process avoids the need to screen libraries  
42 and is useful in finding intron/exon junctions.

43  
44 One PCR method replaces methods which use labeled probes to screen plasmid  
45 libraries and allow one researcher to process only about 3-5 genes in 14-40 days. In the  
46 first step, which can be performed in about two days, any two of a plurality of primers are

designed and synthesized based on a known partial sequence. In step 2, which takes about six to eight hours, the sequence is extended by PCR amplification of a selected library. Steps 3 and 4, which take about one day, are purification of the amplified cDNA and its ligation into an appropriate vector. Step 5, which takes about one day, involves transforming and growing up host bacteria. In step 6, which takes approximately five hours, PCR is used to screen bacterial clones for extended sequence. The final steps, which take about one day, involve the preparation and sequencing of selected clones.

If the full length cDNA has not been obtained, the entire procedure is repeated using either the original library or some other preferred library. The preferred library may be one that has been size-selected to include only larger cDNAs or may consist of single or combined commercially available libraries, e.g., from Clontech Laboratories, Inc. (Palo Alto, Calif.). The cDNA library may have been prepared with oligo (dT) or random priming. Random primed libraries are preferred in that they will contain more sequences which contain 5' ends of genes. A randomly primed library may be particularly useful if an oligo (dT) library does not yield a complete gene. It must be noted that the larger and more complex the protein, the less likely it is that the complete gene will be found in a single plasmid.

CLONTECH PCR-Select.TM. cDNA Subtraction (Clontech Laboratories, Inc., Palo Alto, Calif.) is yet another means by which differentially expressed genes may be isolated. The procedure allows for the isolation of transcripts present in one mRNA population which is absent, or found in reduced numbers, in a second population of mRNA. Rare transcripts may be enriched 1000-fold.

Another method for analyzing either the size or the nucleotide sequence of PCR products is capillary electrophoresis. Systems for rapid sequencing are available from Perkin Elmer (Foster City Calif.), Beckman Instruments (Fullerton, Calif.), and other companies. Capillary sequencing employs flowable polymers for electrophoretic separation, four different fluorescent dyes (one for each nucleotide) which are laser activated, and detection of the emitted wavelengths by a charge coupled device camera. Output/light intensity is converted to electrical signal using appropriate software (eg. Genotyper.TM. and Sequence Navigators.TM. from Perkin Elmer) and the entire process from loading of samples to computer analysis and electronic data display is computer controlled. Capillary electrophoresis provides greater resolution and is many times faster than standard gel based procedures. It is particularly suited to the sequencing of small pieces of DNA which might be present in limited amounts in a particular sample. The reproducible sequencing of up to 350 bp of M13 phage DNA in 30 min has been reported (Ruiz-Martinez M. C. et al. [1993] Anal Chem 65:2851-8).

Polynucleotides and proteins. Polynucleotides of the subject invention can be defined according to several parameters. One characteristic is the biological activity of the protein products as identified herein. The proteins and genes of the subject invention can be further defined by their amino acid and nucleotide sequences. The sequences of the molecules can be defined in terms of homology to certain exemplified sequences as well as in terms of the ability to hybridize with, or be amplified by, certain exemplified probes



1 and primers. Additional primers and probes can readily be constructed by those skilled in  
2 the art such that alternate polynucleotide sequences encoding the same amino acid  
3 sequences can be used to identify and/or characterize additional genes. The proteins of  
4 the subject invention can also be identified based on their immunoreactivity with certain  
5 antibodies.

6  
7 The polynucleotides and proteins of the subject invention include portions, fragments,  
8 variants, and mutants of the full-length sequences as well as fusions and chimerics, so  
9 long as the encoded protein retains the characteristic biological activity of the proteins  
10 identified herein. As used herein, the terms "variants" or "variations" of genes refer to  
11 nucleotide sequences that encode the same proteins or which encode equivalent  
12 proteins having equivalent biological activity. As used herein, the term "equivalent  
13 proteins" refers to proteins having the same or essentially the same biological activity as  
14 the exemplified proteins.

15  
16 Variations of genes may be readily constructed using standard techniques such as site-  
17 directed mutagenesis and other methods of making point mutations and by DNA  
18 shuffling, for example. In addition, gene and protein fragments can be made using  
19 commercially available exonucleases, endonucleases, and proteases according to  
20 standard procedures. For example, enzymes such as Bal31 can be used to  
21 systematically cut off nucleotides from the ends of genes. Also, genes that encode  
22 fragments may be obtained using a variety of restriction enzymes. Proteases may be  
23 used to directly obtain active fragments of these proteins. Of course, molecular  
24 techniques for cloning polynucleotides and producing gene constructs of interest are also  
25 well known in the art. In vitro evaluation techniques, such as MAXYGEN's "Molecular  
26 Breeding" can also be applied to practice the subject invention.

27  
28 Because of the redundancy of the genetic code, a variety of different DNA sequences  
29 can encode the amino acid sequences encoded by the polynucleotide sequences  
30 disclosed herein. It is well within the skill of a person trained in the art to create these  
31 alternative DNA sequences encoding proteins having the same, or essentially the same,  
32 amino acid sequence. These variant DNA sequences are within the scope of the subject  
33 invention. As used herein, reference to "essentially the same" sequence refers to  
34 sequences that have amino acid substitutions, deletions, additions, or insertions that do  
35 not materially affect biological activity. Fragments retaining the characteristic biological  
36 activity are also included in this definition.

37  
38 A farther method for identifying genes and polynucleotides (and the proteins encoded  
39 thereby) of the subject invention is through the use of oligonucleotide probes. Probes  
40 provide a rapid method for identifying genes of the subject invention. The nucleotide  
41 segments that are used as probes according to the invention can be synthesized using a  
42 DNA synthesizer and standard procedures.

43  
44 The subject invention comprises variant or equivalent proteins (and nucleotide  
45 sequences coding for equivalent proteins) having the same or similar biological activity of  
46 proteins encoded by the exemplified polynucleotides. Equivalent proteins will have amino

1 acid similarity with an exemplified protein (or peptide). The amino acid identity will  
2 typically be greater than 60%. Preferably, the amino acid identity will be greater than  
3 75%. More preferably, the amino acid identity will be greater than 80%, and even more  
4 preferably greater than 90%. Most preferably, amino acid identity will be greater than  
5 95%. (Likewise, the polynucleotides that encode the subject polypeptides will also have  
6 corresponding identities in these preferred ranges.) These identities are as determined  
7 using standard alignment techniques for determining amino acid identity. The amino acid  
8 identity/similarity/homology will be highest in critical regions of the protein including those  
9 regions that account for biological activity or that are involved in the determination of  
10 three-dimensional configuration that is ultimately responsible for the biological activity. In  
11 this regard, certain amino acid substitutions are acceptable and can be expected if these  
12 substitutions are in regions which are not critical to activity or are conservative amino  
13 acid substitutions which do not affect the three-dimensional configuration of the  
14 molecule. For example, amino acids may be placed in the following classes: non-polar,  
15 uncharged polar, basic, and acidic. Conservative substitutions whereby an amino acid of  
16 one class is replaced with another amino acid of the same type fall within the scope of  
17 the subject invention so long as the substitution does not materially alter the biological  
18 activity of the compound. Below is a list of examples of amino acids belonging to various  
19 classes.

20  
21 Class of Amino Acid Examples of Amino Acids Nonpolar Ala, Val, Leu, Ile, Pro, Met, Phe,  
22 Trp Uncharged Polar Gly, Ser, Thr, Cys, Tyr, Asn, Gln Acidic Asp, Glu Basic Lys, His  
23

24 In some instances, non-conservative substitutions can also be made.  
25

26 As used herein, reference to "isolated" polynucleotides and/or "purified" proteins refers to  
27 these molecules when they are not associated with the other molecules with which they  
28 would be found in nature. Thus, reference to "isolated" and/or "purified" signifies the  
29 involvement of the "hand of man" as described herein. Reference to "heterologous"  
30 proteins, genes, and gene constructs, also signifies the involvement of the "hand of  
31 man."  
32

### 33 Vectors/Host Cells 34

35 The invention also provides vectors containing the nucleic acid molecules described  
36 herein. The term "vector" refers to a vehicle, preferably a nucleic acid molecule, which  
37 can transport the nucleic acid molecules. When the vector is a nucleic acid molecule, the  
38 nucleic acid molecules are covalently linked to the vector nucleic acid. With this aspect of  
39 the invention, the vector includes a plasmid, single or double stranded phage, a single or  
40 double stranded RNA or DNA viral vector, or artificial chromosome, such as a BAC, PAC,  
41 YAC, OR MAC.  
42

43 A vector can be maintained in the host cell as an extrachromosomal element where it  
44 replicates and produces additional copies of the nucleic acid molecules. Alternatively, the  
45 vector may integrate into the host cell genome and produce additional copies of the  
46 nucleic acid molecules when the host cell replicates.

1  
2 The invention provides vectors for the maintenance (cloning vectors) or vectors for  
3 expression (expression vectors) of the nucleic acid molecules. The vectors can function  
4 in prokaryotic or eukaryotic cells or in both (shuttle vectors).

5  
6 Expression vectors contain cis-acting regulatory regions that are operably linked in the  
7 vector to the nucleic acid molecules such that transcription of the nucleic acid molecules  
8 is allowed in a host cell. The nucleic acid molecules can be introduced into the host cell  
9 with a separate nucleic acid molecule capable of affecting transcription. Thus, the second  
10 nucleic acid molecule may provide a trans-acting factor interacting with the cis-regulatory  
11 control region to allow transcription of the nucleic acid molecules from the vector.  
12 Alternatively, a trans-acting factor may be supplied by the host cell. Finally, a trans-acting  
13 factor can be produced from the vector itself. It is understood, however, that in some  
14 embodiments, transcription and/or translation of the nucleic acid molecules can occur in  
15 a cell-free system.

16  
17 The regulatory sequence to which the nucleic acid molecules described herein can be  
18 operably linked include promoters for directing mRNA transcription. These include, but  
19 are not limited to, the left promoter from bacteriophage .lambda., the lac, TRP, and TAC  
20 promoters from E. Coli, the early and late promoters from SV40, the CMV immediate  
21 early promoter, the adenovirus early and late promoters, and retrovirus long-terminal  
22 repeats.

23  
24 In addition to control regions that promote transcription, expression vectors may also  
25 include regions that modulate transcription, such as repressor binding sites and  
26 enhancers. Examples include the SV40 enhancer, the cytomegalovirus immediate early  
27 enhancer, polyoma enhancer, adenovirus enhancers, and retrovirus LTR enhancers.

28  
29 In addition to containing sites for transcription initiation and control, expression vectors  
30 can also contain sequences necessary for transcription termination and, in the  
31 transcribed region a ribosome binding site for translation. Other regulatory control  
32 elements for expression include initiation and termination codons as well as  
33 polyadenylation signals. The person of ordinary skill in the art would be aware of the  
34 numerous regulatory sequences that are useful in expression vectors. Such regulatory  
35 sequences are described, for example, in Sambrook et al., Molecular Cloning: A  
36 Laboratory Manual. 2nd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor,  
37 N.Y., (1989).

38  
39 A variety of expression vectors can be used to express a nucleic acid molecule. Such  
40 vectors include chromosomal, episomal, and virus-derived vectors, for example vectors  
41 derived from bacterial plasmids, from bacteriophage, from yeast episomes, from yeast  
42 chromosomal elements, including yeast artificial chromosomes, from viruses such as  
43 baculoviruses, papovaviruses such as SV40, Vaccinia viruses, adenoviruses, poxviruses,  
44 pseudorabies viruses, and retroviruses. Vectors may also be derived from combinations  
45 of these sources such as those derived from plasmid and bacteriophage genetic  
46 elements, e.g. cosmids and phagemids. Appropriate cloning and expression vectors for

1 prokaryotic and eukaryotic hosts are described in Sambrook et al., Molecular Cloning: A  
2 Laboratory Manual. 2nd. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor,  
3 N.Y., (1989).

5 The regulatory sequence may provide constitutive expression in one or more host cells  
6 (i.e. tissue specific) or may provide for inducible expression in one or more cell types  
7 such as by temperature, nutrient additive, or exogenous factor such as a hormone or  
8 other ligand. A variety of vectors providing for constitutive and inducible expression in  
9 prokaryotic and eukaryotic hosts are well known to those of ordinary skill in the art.

11 The nucleic acid molecules can be inserted into the vector nucleic acid by well-known  
12 methodology. Generally, the DNA sequence that will ultimately be expressed is joined to  
13 an expression vector by cleaving the DNA sequence and the expression vector with one  
14 or more restriction enzymes and then ligating the fragments together. Procedures for  
15 restriction enzyme digestion and ligation are well known to those of ordinary skill in the  
16 art.

18 The vector containing the appropriate nucleic acid molecule can be introduced into an  
19 appropriate host cell for propagation or expression using well-known techniques.  
20 Bacterial cells include, but are not limited to, E. coli, Streptomyces, and Salmonella  
21 typhimurium. Eukaryotic cells include, but are not limited to, yeast, insect cells such as  
22 Drosophila, animal cells such as COS and CHO cells, and plant cells.

24 As described herein, it may be desirable to express the peptide as a fusion protein.  
25 Accordingly, the invention provides fusion vectors that allow for the production of the  
26 peptides. Fusion vectors can increase the expression of a recombinant protein, increase  
27 the solubility of the recombinant protein, and aid in the purification of the protein by acting  
28 for example as a ligand for affinity purification. A proteolytic cleavage site may be  
29 introduced at the junction of the fusion moiety so that the desired peptide can ultimately  
30 be separated from the fusion moiety. Proteolytic enzymes include, but are not limited to,  
31 factor Xa, thrombin, and enteroenzyme. Typical fusion expression vectors include pGEX  
32 (Smith et al., Gene 67:31-40 (1988)), pMAL (New England Biolabs, Beverly, Mass.) and  
33 pRIT5 (Pharmacia, Piscataway, N.J.) which fuse glutathione S-transferase (GST),  
34 maltose E binding protein, or protein A, respectively, to the target recombinant protein.  
35 Examples of suitable inducible non-fusion E. coli expression vectors include pTrc  
36 (Amann et al, Gene 69:301-315 (1988)) and pET 11d (Studier et al, Gene Expression  
37 Technology: Methods in Enzymology 185:60-89 (1990)).

39 Recombinant protein expression can be maximized in host bacteria by providing a  
40 genetic background wherein the host cell has an impaired capacity to proteolytically  
41 cleave the recombinant protein. (Gottesman, S., Gene Expression Technology: Methods  
42 in Enzymology 185, Academic Press, San Diego, Calif. (1990) 119-128). Alternatively,  
43 the sequence of the nucleic acid molecule of interest can be altered to provide  
44 preferential codon usage for a specific host cell, for example E. coli. (Wada et al., Nucleic  
45 Acids Res. 20:2111-2118 (1992)).

1 The nucleic acid molecules can also be expressed by expression vectors that are  
2 operative in yeast. Examples of vectors for expression in yeast e.g., *S. cerevisiae* include  
3 pYepSecI (Baldari, et al, EMBO J. 6:229-234 (1987)), pMFa (Kujan et al, Cell 30:933-  
4 943(1982)), pJRY88 (Schultz et al, Gene 54:113-123 (1987)), and pYES2 (Invitrogen  
5 Corporation, San Diego, Calif.).

6  
7 The nucleic acid molecules can also be expressed in insect cells using, for example,  
8 baculovirus expression vectors. Baculovirus vectors available for expression of proteins  
9 in cultured insect cells (e.g., Sf9 cells) include the pAc series (Smith et al, Mol. Cell Biol.  
10 3:2156-2165 (1983)) and the pVL series (Lucklow et al, Virology 170:31-39 (1989)).

11  
12 In certain embodiments of the invention, the nucleic acid molecules described herein are  
13 expressed in mammalian cells using mammalian expression vectors. Examples of  
14 mammalian expression vectors include pCDM8 (Seed, B. Nature 329:840(1987)) and  
15 pMT2PC (Kaufman et al., EMBO J. 6:187-195 (1987)).

16  
17 The expression vectors listed herein are provided by way of example only of the well-  
18 known vectors available to those of ordinary skill in the art that would be useful to  
19 express the nucleic acid molecules. The person of ordinary skill in the art would be aware  
20 of other vectors suitable for maintenance propagation or expression of the nucleic acid  
21 molecules described herein. These are found for example in Sambrook, J., Fritsh, E. F.,  
22 and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor  
23 Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

24  
25 The invention also encompasses vectors in which the nucleic acid sequences described  
26 herein are cloned into the vector in reverse orientation, but operably linked to a  
27 regulatory sequence that permits transcription of antisense RNA. Thus, an antisense  
28 transcript can be produced to all, or to a portion, of the nucleic acid molecule sequences  
29 described herein, including both coding and non-coding regions. Expression of this  
30 antisense RNA is subject to each of the parameters described above in relation to  
31 expression of the sense RNA (regulatory sequences, constitutive or inducible expression,  
32 tissue-specific expression).

33  
34 The invention also relates to recombinant host cells containing the vectors described  
35 herein. Host cells therefore include prokaryotic cells, lower eukaryotic cells such as  
36 yeast, other eukaryotic cells such as insect cells, and higher eukaryotic cells such as  
37 mammalian cells.

38  
39 The recombinant host cells are prepared by introducing the vector constructs described  
40 herein into the cells by techniques readily available to the person of ordinary skill in the  
41 art. These include, but are not limited to, calcium phosphate transfection, DEAE-dextran-  
42 mediated transfection, cationic lipid-mediated transfection, electroporation, transduction,  
43 infection, lipofection, and other techniques such as those found in Sambrook, et al.  
44 (Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold  
45 Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).



1 Host cells can contain more than one vector. Thus, different nucleotide sequences can  
2 be introduced on different vectors of the same cell. Similarly, the nucleic acid molecules  
3 can be introduced either alone or with other nucleic acid molecules that are not related to  
4 the nucleic acid molecules such as those providing trans-acting factors for expression  
5 vectors. When more than one vector is introduced into a cell, the vectors can be  
6 introduced independently, co-introduced or joined to the nucleic acid molecule vector.

8 In the case of bacteriophage and viral vectors, these can be introduced into cells as  
9 packaged or encapsulated virus by standard procedures for infection and transduction.  
10 Viral vectors can be replication-competent or replication-defective. In the case in which  
11 viral replication is defective, replication will occur in host cells providing functions that  
12 complement the defects.

14 Vectors generally include selectable markers that enable the selection of the  
15 subpopulation of cells that contain the recombinant vector constructs. The marker can be  
16 contained in the same vector that contains the nucleic acid molecules described herein  
17 or may be on a separate vector. Markers include tetracycline or ampicillin-resistance  
18 genes for prokaryotic host cells and dihydrofolate reductase or neomycin resistance for  
19 eukaryotic host cells. However, any marker that provides selection for a phenotypic trait  
20 will be effective.

22 While the mature proteins can be produced in bacteria, yeast, mammalian cells, and  
23 other cells under the control of the appropriate regulatory sequences, cell- free  
24 transcription and translation systems can also be used to produce these proteins using  
25 RNA derived from the DNA constructs described herein.

27 Where secretion of the peptide is desired, which is difficult to achieve with multi-  
28 transmembrane domain containing proteins such as MHC Class I-binding peptides,  
29 appropriate secretion signals are incorporated into the vector. The signal sequence can  
30 be endogenous to the peptides or heterologous to these peptides.

32 The expressed protein can be isolated from the host cell by standard disruption  
33 procedures, including freeze thaw, sonication, mechanical disruption, use of lysing  
34 agents and the like. The peptide can then be recovered and purified by well-known  
35 purification methods including ammonium sulfate precipitation, acid extraction, anion or  
36 cationic exchange chromatography, phosphocellulose chromatography, hydrophobic-  
37 interaction chromatography, affinity chromatography, hydroxylapatite chromatography,  
38 lectin chromatography, or high performance liquid chromatography.

40 It is also understood that depending upon the host cell in recombinant production of the  
41 peptides described herein, the peptides can have various glycosylation patterns,  
42 depending upon the cell, or maybe non-glycosylated as when produced in bacteria. In  
43 addition, the peptides may include an initial modified methionine in some cases as a  
44 result of a host-mediated process.

46 There are many methods for introducing a heterologous gene or polynucleotide into a

1 host cell or cells under conditions that allow for stable maintenance and expression of the  
2 gene or polynucleotide. These methods are well known to those skilled in the art.  
3 Synthetic genes, such as, for example, those genes modified to enhance expression in a  
4 heterologous host (such as by preferred codon usage or by the use of adjoining,  
5 downstream, or upstream enhancers) that are functionally equivalent to the genes (and  
6 which encode equivalent proteins) can also be used to transfect hosts. Methods for the  
7 production of synthetic genes are known in the art. Recombinant hosts can be used for  
8 the expression or propagation of genes and polynucleotides of the present invention. The  
9 genes and polynucleotides within the scope of the present invention can be introduced  
10 into a wide variety of microbial or plant hosts, such as bacterial cells, yeast, insect cells,  
11 plant cell cultures or plants, mammalian cells. For example, T. parva nucleic acids can  
12 be expressed in a recombinant baculovirus (BV) possessing an optimized promoter and  
13 translation initiation region operably linked to the T. parva nucleic acid. In a preferred  
14 embodiment, an optimized promoter and translation initiation region are operably linked  
15 to the T. parva nucleic acid. In one embodiment, insect host cells can be transformed  
16 with baculovirus expressing T. parva nucleic acids. In still another embodiment, Tn5  
17 (Trichoplusia ni or High Five.TM.) host cells can be transformed with baculovirus  
18 expressing T. parva nucleic acids.

19  
20 T. parva "recombinant protein", is a protein made using recombinant techniques, i.e.  
21 through the expression of a recombinant nucleic acid as depicted above. A recombinant  
22 protein is distinguished from naturally occurring protein by at least one or more  
23 characteristics. For example, the protein may be isolated or purified away from some or  
24 all of the proteins and compounds with which it is normally associated in its wild type  
25 host, and thus may be substantially pure. For example, an isolated protein is  
26 unaccompanied by at least some of the material with which it is normally associated in its  
27 natural state, preferably constituting at least about 0.5%, more preferably at least about  
28 5% by weight of the total protein in a given sample. A substantially pure protein  
29 comprises at least about 75% by weight of the total protein, with at least about 80%  
30 being preferred, and at least about 90% being particularly preferred. The definition  
31 includes the production of a protein from one organism in a different organism or host  
32 cell. Alternatively, the protein may be made at a significantly higher concentration than is  
33 normally seen, through the use of an inducible promoter or high expression promoter,  
34 such that the protein is made at increased concentration levels. Alternatively, the protein  
35 may be in a form not normally found in nature, as in the addition of an epitope tag or  
36 amino acid substitutions, insertions and/or deletions, as discussed below.

37  
38 Included in the definition of T. parva antigen polypeptides are T. parva polypeptide  
39 variants. These variants fall into one or more of three classes: substitutional, insertional  
40 or deletional variants. These variants ordinarily are prepared by site specific mutagenesis  
41 of nucleotides in the DNA encoding a T. parva antigen polypeptide, using cassette or  
42 PCR mutagenesis, scanning mutagenesis, gene shuffling or other techniques well known  
43 in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in  
44 recombinant cell culture as outlined above. However, variant T. parva polypeptide  
45 fragments having up to about 100-150 residues may be prepared by in vitro synthesis  
46 using established techniques. Amino acid sequence variants are characterized by the  
47 predetermined nature of the variation, a feature that sets them apart from naturally

1 occurring allelic or interspecies variation of the T. parva antigen polypeptide amino acid  
2 sequence.

3  
4 While the site or region for introducing an amino acid sequence variation is  
5 predetermined, the mutation per se need not be predetermined. For example, in order to  
6 optimize the performance of a mutation at a given site, random mutagenesis may be  
7 conducted at the target codon or region and the expressed T. parva antigen polypeptide  
8 variants can be screened for the optimal combination of desired activity. Techniques for  
9 making mutations at predetermined sites in DNA having a known sequence are well  
10 known. For example, the variations can be made using oligonucleotide-mediated site-  
11 directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl.  
12 Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)],  
13 restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA,  
14 317:415 (1986)], all of which are expressly incorporated by reference, PCR mutagenesis,  
15 or other known techniques can be performed on the cloned DNA to produce T. parva  
16 antigen polypeptide variant DNA. Scanning amino acid analysis can also be employed to  
17 identify one or more amino acids along a contiguous sequence. Among the preferred  
18 scanning amino acids are relatively small, neutral amino acids. Such amino acids include  
19 alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid  
20 among this group because it eliminates the side-chain beyond the beta-carbon and is less  
21 likely to alter the main-chain conformation of the variant [Cunningham and Wells,  
22 Science, 244: 1081-1085 (1989), which is expressly incorporated by reference]. Alanine is  
23 also typically preferred because it is the most common amino acid. Further, it is frequently  
24 found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman &  
25 Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976), which are expressly incorporated by  
26 reference]. If alanine substitution does not yield adequate amounts of variant, an isoteric  
27 amino acid can be used. Screening of the mutants or variants is done using Elispot and/or  
28 bioassays of T. parva antigen polypeptide activities and/or properties as described herein.

29  
30  
31 The present invention further provides fragments of the antigen peptides, in addition to  
32 proteins and peptides that comprise and consist of such fragments, particularly those  
33 comprising the residues identified in FIG. 2. The fragments to which the invention  
34 pertains, however, are not to be construed as encompassing fragments that may be  
35 disclosed publicly prior to the present invention.

36  
37 As used herein, a fragment comprises at least 8, 10, 12, 14, 16, or more contiguous  
38 amino acid residues from an antigen peptide. Such fragments can be chosen based on  
39 the ability to retain one or more of the biological activities of the antigen peptide or could  
40 be chosen for the ability to perform a function, e.g. bind a substrate or act as an  
41 immunogen. Particularly important fragments are biologically active fragments, peptides  
42 that are, for example, about 8 or more amino acids in length. Such fragments will typically  
43 comprise a domain or motif of the antigen peptide, e.g., active site, a transmembrane  
44 domain or a substrate-binding domain. Further, possible fragments include, but are not  
45 limited to, domain or motif containing fragments, soluble peptide fragments, and  
46 fragments containing immunogenic structures. Predicted domains and functional sites are  
47 readily identifiable by computer programs well known and readily available to those of skill

1 in the art (e.g., PROSITE analysis). The results of one such analysis are provided in FIG.  
2 2.

3  
4 Polypeptides often contain amino acids other than the 20 amino acids commonly referred  
5 to as the 20 naturally occurring amino acids. Further, many amino acids, including the  
6 terminal amino acids, may be modified by natural processes, such as processing and  
7 other post-translational modifications, or by chemical modification techniques well known  
8 in the art. Common modifications that occur naturally in antigen peptides are described in  
9 basic texts, detailed monographs, and the research literature, and they are well known to  
10 those of skill in the art.

11  
12 Known modifications include, but are not limited to, acetylation, acylation, ADP-  
13 ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme  
14 moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment  
15 of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking,  
16 cyclization, disulfide bond formation, demethylation, formation of covalent crosslinks,  
17 formation of cystine, formation of pyroglutamate, formylation, gamma carboxylation,  
18 glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation,  
19 oxidation, proteolytic processing, phosphorylation, prenylation, racemization,  
20 selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such  
21 as arginylation, and ubiquitination.

22  
23 Such modifications are well known to those of skill in the art and have been described in  
24 great detail in the scientific literature. Several particularly common modifications,  
25 glycosylation, lipid attachment, sulfation, gamma-carboxylation of glutamic acid residues,  
26 hydroxylation and ADP-ribosylation, for instance, are described in most basic texts, such  
27 as *Proteins--Structure and Molecular Properties*, 2<sup>nd</sup> Ed., T. E. Creighton, W. H.  
28 Freeman and Company, New York (1993). Many detailed reviews are available on this  
29 subject, such as by Wold, F., *Posttranslational Covalent Modification of Proteins*, B. C.  
30 Johnson, Ed., Academic Press, New York 1-12 (1983); Seifter et al. (*Meth. Enzymol.* 182:  
31 626-646 (1990)) and Rattan et al. (*Ann. N.Y Acad. Sci.* 663:48-62 (1992)).

32  
33 Accordingly, the antigen peptides of the present invention also encompass derivatives or  
34 analogs in which a substituted amino acid residue is not one encoded by the genetic  
35 code, in which a substituent group is included, in which the mature antigen peptide is  
36 fused with another compound, such as a compound to increase the half-life of the antigen  
37 peptide (for example, polyethylene glycol), or in which the additional amino acids are  
38 fused to the mature antigen peptide, such as a leader or secretory sequence or a  
39 sequence for purification of the mature antigen peptide or a pro-protein sequence.

#### 40 41 Antibody Production

42  
43 Although an amino acid sequence or oligopeptide used for antibody induction does not  
44 require biological activity, it must be immunogenic. A peptide, polypeptide, or protein used  
45 to induce specific antibodies may have an amino acid sequence consisting of at least five  
46 amino acids and preferably at least 10 amino acids. Short stretches of amino acid

1 sequence may be genetically or chemically fused with those of another protein such as  
2 keyhole limpet hemocyanin, and the chimeric peptide used for antibody production.  
3 Alternatively, the oligopeptide may be of sufficient length to contain an entire domain.

4  
5 Antibodies specific for peptides, polypeptides, or proteins may be produced by inoculation  
6 of an appropriate animal with an antigenic fragment of the peptide, polypeptide, or  
7 protein. Antibody production includes not only the stimulation of an immune response by  
8 injection into animals, but also analogous processes such as the production of synthetic  
9 antibodies, the screening of recombinant immunoglobulin libraries for specific-binding  
10 molecules (Orlandi R. et al. [1989] PNAS 86:3833-3837, or Huse W. D. et al. [1989]  
11 Science 256:1275-1281), or the in vitro stimulation of lymphocyte populations. Current  
12 technology (Winter G. and Milstein C. [1991] Nature 349:293-299) provides for a number  
13 of highly specific binding reagents based on the principles of antibody formation. These  
14 techniques may be adapted to produce molecules which specifically bind antigen  
15 peptides. Antibodies or other appropriate molecules generated against a specific  
16 immunogenic peptide fragment or oligopeptide can be used in Western analysis, enzyme-  
17 linked immunosorbent assays (ELISA) or similar tests to establish the presence of or to  
18 quantitate amounts of peptide, polypeptide, or protein in normal, diseased, or transformed  
19 cells, tissues, organs, or organisms as well as liquid suspensions containing said peptide,  
20 polypeptide, or protein.

21  
22 The invention also provides antibodies that selectively bind to one of the peptides of the  
23 present invention, a protein comprising such a peptide, as well as variants and fragments  
24 thereof. As used herein, an antibody selectively binds a target peptide when it binds the  
25 target peptide and does not significantly bind to unrelated proteins. An antibody is still  
26 considered to selectively bind a peptide even if it also binds to other proteins that are not  
27 substantially homologous with the target peptide so long as such proteins share  
28 homology with a fragment or domain of the peptide target of the antibody. In this case, it  
29 would be understood that antibody binding to the peptide is still selective despite some  
30 degree of cross-reactivity.

31  
32 As used herein, an antibody is defined in terms consistent with that recognized within the  
33 art: they are multi-subunit proteins produced by a mammalian organism in response to an  
34 antigen challenge. The antibodies of the present invention include polyclonal antibodies  
35 and monoclonal antibodies, as well as fragments of such antibodies, including, but not  
36 limited to, Fab or F(ab').sub.2, and Fv fragments.

37  
38 Many methods are known for generating and/or identifying antibodies to a given target  
39 peptide. Several such methods are described by Harlow, Antibodies, Cold Spring Harbor  
40 Press, (1989).

41  
42 In general, to generate antibodies, an isolated peptide is used as an immunogen and is  
43 administered to a mammalian organism, such as a rat, rabbit or mouse. The full-length  
44 protein, an antigenic peptide fragment or a fusion protein can be used. Particularly  
45 important fragments are those covering functional domains, such as the domains  
46 identified in the figures, and domain of sequence homology or divergence amongst the

1 family, such as those that can readily be identified using protein alignment methods and  
2 as presented in the Figures.

3  
4 Antibodies are preferably prepared from regions or discrete fragments of the antigen  
5 proteins. Antibodies can be prepared from any region of the peptide as described herein.  
6 However, preferred regions will include those involved in function/activity and/or  
7 antigen/binding partner interaction.

8  
9 An antigenic fragment will typically comprise at least 8 contiguous amino acid residues.  
10 The antigenic peptide can comprise, however, at least 10, 12, 14, 16 or more amino acid  
11 residues. Such fragments can be selected on a physical property, such as fragments  
12 correspond to regions that are located on the surface of the protein, e.g., hydrophilic  
13 regions or can be selected based on sequence uniqueness.

14  
15 Detection on an antibody of the present invention can be facilitated by coupling (i.e.,  
16 physically linking) the antibody to a detectable substance. Examples of detectable  
17 substances include various enzymes, prosthetic groups, fluorescent materials,  
18 luminescent materials, bioluminescent materials, and radioactive materials. Examples of  
19 suitable enzymes include horseradish peroxidase, alkaline phosphatase, .beta.-  
20 galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes  
21 include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials  
22 include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine,  
23 dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a  
24 luminescent material includes luminol; examples of bioluminescent materials include  
25 luciferase, luciferin, and aequorin, and examples of suitable radioactive material include  
26 .sup.125I, .sup.131I, .sup.35S or .sup.3H.

27  
28 The antibodies can be used to isolate one of the proteins of the present invention by  
29 standard techniques, such as affinity chromatography or immunoprecipitation. The  
30 antibodies can facilitate the purification of the natural protein from cells and recombinantly  
31 produced protein expressed in host cells. In addition, such antibodies are useful to detect  
32 the presence of one of the proteins of the present invention in cells or tissues to  
33 determine the pattern of expression of the protein among various tissues in an organism  
34 and over the course of schizont development. Such antibodies can be used to detect  
35 protein in situ, in vitro, or in a cell lysate or supernatant in order to evaluate the  
36 abundance and pattern of expression. Also, such antibodies can be used to assess  
37 abnormal tissue distribution or abnormal expression during development or progression of  
38 a protozoan infection. Antibody detection of circulating fragments of the full length protein  
39 can be used to identify turnover.

40  
41 Further, the antibodies can be used to assess expression in disease states such as in  
42 active stages of the disease or in lymphocytes harvested from infected animals.  
43 Experimental data as provided in FIG. 22 indicates expression in bovine lymphocytes,  
44 infected with T. parva. If a disease is characterized by a specific mutation in the protein,  
45 antibodies specific for this mutant protein can be used to assay for the presence of the  
46 specific mutant protein.

1  
2 The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a  
3 treatment modality. Accordingly, where treatment is ultimately aimed at preventing or  
4 controlling infection, antibodies directed against the protein or relevant fragments can be  
5 used to monitor therapeutic efficacy.

6  
7 The antibodies are also useful for inhibiting protein function, for example, blocking the  
8 binding of the antigenic peptide to a binding partner such as a substrate. These uses can  
9 also be applied in a therapeutic context in which treatment involves inhibiting the protein's  
10 function. An antibody can be used, for example, to block binding, thus modulating  
11 (agonizing or antagonizing) the peptides activity. Antibodies can be prepared against  
12 specific fragments containing sites required for function or against intact protein that is  
13 associated with a cell or cell membrane. See FIGS. 16-20 or 23, for structural information  
14 relating to the proteins of the present invention.

15  
16 The invention also encompasses kits for using antibodies to detect the presence of a  
17 protein in a biological sample. The kit can comprise antibodies such as a labeled or  
18 labelable antibody and a compound or agent for detecting protein in a biological sample;  
19 means for determining the amount of protein in the sample; means for comparing the  
20 amount of protein in the sample with a standard; and instructions for use. Such a kit can  
21 be supplied to detect a single protein or epitope or can be configured to detect one of a  
22 multitude of epitopes, such as in an antibody detection array. Arrays are described in  
23 detail below for nucleic acid arrays and similar methods have been developed for antibody  
24 arrays.

## 25 26 Nucleic Acid Molecules

27  
28 The present invention further provides isolated nucleic acid molecules that encode a T.  
29 parva antigen peptide or protein of the present invention (cDNA, transcript and genomic  
30 sequence). Such nucleic acid molecules will consist of, consist essentially of, or comprise  
31 a nucleotide sequence that encodes one of the enzyme peptides of the present invention,  
32 an allelic variant thereof, or an ortholog or paralog thereof.

33  
34 As used herein, an "isolated" nucleic acid molecule is one that is separated from other  
35 nucleic acid present in the natural source of the nucleic acid. Preferably, an "isolated"  
36 nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences  
37 located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from  
38 which the nucleic acid is derived. However, there can be some flanking nucleotide  
39 sequences, for example up to about 5KB, 4KB, 3KB, 2KB, or 1KB or less, particularly  
40 contiguous peptide encoding sequences and peptide encoding sequences within the  
41 same gene but separated by introns in the genomic sequence. The important point is that  
42 the nucleic acid is isolated from remote and unimportant flanking sequences such that it  
43 can be subjected to the specific manipulations described herein such as recombinant  
44 expression, preparation of probes and primers, and other uses specific to the nucleic acid  
45 sequences.



1 Moreover, an "isolated" nucleic acid molecule, such as a transcript/cDNA molecule, can  
2 be substantially free of other cellular material, or culture medium when produced by  
3 recombinant techniques, or chemical precursors or other chemicals when chemically  
4 synthesized. However, the nucleic acid molecule can be fused to other coding or  
5 regulatory sequences and still be considered isolated.  
6  
7 For example, recombinant DNA molecules contained in a vector are considered isolated.  
8 Further examples of isolated DNA molecules include recombinant DNA molecules  
9 maintained in heterologous host cells or purified (partially or substantially) DNA molecules  
10 in solution. Isolated RNA molecules include in vivo or in vitro RNA transcripts of the  
11 isolated DNA molecules of the present invention. Isolated nucleic acid molecules  
12 according to the present invention further include such molecules produced synthetically.  
13  
14 Following are examples which illustrate procedures for practicing the invention. These  
15 examples should not be construed as limiting. All percentages are by weight and all  
16 solvent mixture proportions are by volume unless otherwise noted. As used herein and  
17 further defined below, "nucleic acid" may refer to either DNA or RNA, or molecules which  
18 contain both deoxy- and ribonucleotides. The nucleic acids include genomic DNA, cDNA,  
19 and oligonucleotides including sense and anti-sense nucleic acids. Such nucleic acids  
20 may also contain modifications in the ribose-phosphate backbone to increase stability and  
21 half life of such molecules in physiological environments.  
22  
23 The nucleic acid may be double stranded, single stranded, or contain portions of both  
24 double stranded or single stranded sequence. As will be appreciated by those in the art,  
25 the depiction of a single strand ("Watson") also defines the sequence of the other strand  
26 ("Crick"); thus the sequences depicted in the figures also include the complement of the  
27 sequence.  
28  
29 One method for controlling gene expression according to the subject invention provides  
30 materials that would inform the production of double-stranded interfering RNA (dsRNAi),  
31 or RNA-mediated interference (RNAi), using published methods. The terms dsRNAi and  
32 RNAi are used interchangeably herein unless otherwise noted. In a more preferred  
33 embodiment, compositions are useful for regulation of gene expression in bovine  
34 lymphocytes. Such RNA would contain a nucleotide sequence identical to a fragment of  
35 the target gene; however, RNA sequences with insertions, deletions, and point mutations  
36 relative to the target sequence can also be used for inhibition. Sequence identity may  
37 optimized by sequence comparison and alignment algorithms known in the art (see  
38 Gribskov and Devereux, Sequence Analysis Primer, Stockton Press, 1991, and  
39 references cited therein) and calculating the percent difference between the nucleotide  
40 sequences by, for example, the Smith-Waterman algorithm as implemented in the  
41 BESTFIT software program using default parameters (e.g., University of Wisconsin  
42 Genetic Computing Group). Alternatively, the duplex region of the RNA may be defined  
43 functionally as a nucleotide sequence that is capable of hybridizing with a fragment of the  
44 target gene transcript.  
45

1 As disclosed herein, 100% sequence identity between the RNA and the target gene is not  
2 required to practice the present invention. Thus the invention has the advantage of being  
3 able to tolerate sequence variations that might be expected due to genetic mutation,  
4 strain polymorphism, or evolutionary divergence. RNAi molecules of the subject invention  
5 are not limited to those that are targeted to the full-length polynucleotide or gene. Gene  
6 product can be inhibited with a RNAi molecule that is targeted to a portion or fragment of  
7 the exemplified polynucleotides; high homology (90-95%) or greater identity is also  
8 preferred, but not necessarily essential, for such applications.

#### 11 Protein/Peptide Uses

13 The proteins of the present invention can be used in substantial and specific assays  
14 related to the functional information provided in the Figures; for example, to stimulate  
15 CD8+ cytotoxic T cell lines, to raise antibodies or to elicit another immune response; as a  
16 reagent (including the labeled reagent) in assays designed to quantitatively determine  
17 levels of the protein (or its binding partner or ligand) in biological fluids; and as markers  
18 for tissues in which the corresponding protein is preferentially expressed (either  
19 constitutively or at a particular stage of tissue differentiation or development or in a  
20 disease state). Where the protein binds or potentially binds to another protein or ligand  
21 (such as, for example, in a enzyme-effector protein interaction or enzyme-ligand  
22 interaction), the protein can be used to identify the binding partner/ligand so as to develop  
23 a system to identify inhibitors of the binding interaction. Any or all of these uses are  
24 capable of being developed into reagent grade or kit format for commercialization as  
25 commercial products.

27 Methods for performing the uses listed above are well known to those skilled in the art.  
28 References disclosing such methods include "Molecular Cloning: A Laboratory Manual",  
29 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis  
30 eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques",  
31 Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

33 The potential uses of the peptides of the present invention are based primarily on the  
34 source of the protein as well as the class/action of the protein. For example, antigens  
35 isolated from T. parva and their protozoan orthologs serve as targets for identifying  
36 agents for use in mammalian therapeutic applications, e.g. an animal drug, particularly in  
37 modulating a biological or pathological response in a cell, tissue, or protozoan, that  
38 expresses an immunogenic antigen. Experimental data as provided in FIG. 23 indicates  
39 that the antigens of the present invention are expressed in infected host lymphocytes, as  
40 indicated by flow cytometric analysis. Such uses can readily be determined using the  
41 information provided herein, that which is known in the art, and routine experimentation.

43 The proteins of the present invention are also useful in drug screening assays, in cell-  
44 based or cell-free systems. Cell-based systems can be native, i.e., cells that normally  
45 express the enzyme, as a biopsy or expanded in cell culture. Experimental data as  
46 provided in FIG. 22 indicates expression in infected bovine lymphocytes. In an alternate  
47 embodiment, cell-based assays involve recombinant host cells expressing the antigen

1 protein.

2

3 The polypeptides can be used to identify compounds that modulate pathogen activity as  
4 measured by the antigen in its natural state or an altered form that causes a specific  
5 disease or pathology associated with the protozoan. Both the antigens of the present  
6 invention and appropriate variants and fragments can be used in high-throughput screens  
7 to assay candidate compounds for the ability to bind to the antigen. These compounds  
8 can be further screened against a functional T-cell antigen to determine the effect of the  
9 compound on the enzyme activity. Further, these compounds can be tested in animal or  
10 invertebrate systems to determine activity/effectiveness.

11

12 The invention further includes other end point assays to identify compounds that  
13 modulate (stimulate or inhibit) protozoan antigen expression.

14

15 The proteins of the present invention are also useful in competition binding assays in  
16 methods designed to discover compounds that interact with the T. parva antigens (e.g.  
17 binding partners and/or ligands). Thus, a compound is exposed to an antigen polypeptide  
18 under conditions that allow the compound to bind or to otherwise interact with the  
19 polypeptide. Soluble antigen polypeptide is also added to the mixture. If the test  
20 compound interacts with the soluble antigen polypeptide, it decreases the amount of  
21 complex formed or activity from the antigen target. This type of assay is particularly useful  
22 in cases in which compounds are sought that interact with specific regions of the antigen.  
23 Thus, the soluble polypeptide that competes with the target antigen region is designed to  
24 contain peptide sequences corresponding to the region of interest.

25

26 To perform cell free drug screening assays, it is sometimes desirable to immobilize either  
27 the antigen protein, or fragment, or its target molecule to facilitate separation of  
28 complexes from uncomplexed forms of one or both of the proteins, as well as to  
29 accommodate automation of the assay.

30

31 Techniques for immobilizing proteins on matrices can be used in the drug screening  
32 assays. In one embodiment, a fusion protein can be provided which adds a domain that  
33 allows the protein to be bound to a matrix. For example, glutathione-S-transferase fusion  
34 proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis,  
35 Mo.) or glutathione derivatized microtitre plates, which are then combined with the cell  
36 lysates (e.g., <sup>35</sup>S-labeled) and the candidate compound, and the mixture incubated  
37 under conditions conducive to complex formation (e.g., at physiological conditions for salt  
38 and pH). Following incubation, the beads are washed to remove any unbound label, and  
39 the matrix immobilized and radiolabel determined directly, or in the supernatant after the  
40 complexes are dissociated. Alternatively, the complexes can be dissociated from the  
41 matrix, separated by SDS-PAGE, and the level of antigen-binding protein found in the  
42 bead fraction quantitated from the gel using standard electrophoretic techniques. For  
43 example, either the polypeptide or its target molecule can be immobilized utilizing  
44 conjugation of biotin and streptavidin using techniques well known in the art. Alternatively,  
45 antibodies reactive with the protein but which do not interfere with binding of the protein to  
46 its target molecule can be derivatized to the wells of the plate, and the protein trapped in

1 the wells by antibody conjugation. Preparations of an antigen-binding protein and a  
2 candidate compound are incubated in the antigen protein-presenting wells and the  
3 amount of complex trapped in the well can be quantitated. Methods for detecting such  
4 complexes, in addition to those described above for the GST-immobilized complexes,  
5 include immunodetection of complexes using antibodies reactive with the antigen protein  
6 target molecule, or which are reactive with antigen protein and compete with the target  
7 molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity  
8 associated with the presence of the target molecule.

9  
10 Agents that modulate one of the antigens of the present invention can be identified using  
11 one or more of the above assays, alone or in combination. It is generally preferable to use  
12 a cell-based or cell free system first and then confirm activity in an animal or other model  
13 system. Such model systems are well known in the art and can readily be employed in  
14 this context.

15  
16  
17 In yet another aspect of the invention, the antigen proteins can be used as "bait proteins"  
18 in a two-hybrid assay or three-hybrid assay (see, e.g., U.S. Pat. No. 5,283,317; Zervos et  
19 al. (1993) Cell 72:223-232; Madura et al. (1993) J. Biol. Chem. 268:12046-12054; Bartel  
20 et al. (1993) Biotechniques 14:920-924; Iwabuchi et al. (1993) Oncogene 8:1693-1696;  
21 and Brent WO94/10300), to identify other proteins, which bind to or interact with the  
22 antigen and are involved in T-cell induction activity. Such antigen-binding proteins are  
23 also likely to be involved in the propagation of signals by antigen-binding proteins or  
24 antigen targets as, for example, downstream elements of a MHC Class I-mediated  
25 signaling pathway. Alternatively, such antigen-binding proteins could act in a manner as  
26 to be inhibitors of an immune response.

27  
28  
29 This invention further pertains to novel agents identified by the above-described  
30 screening assays. Accordingly, it is within the scope of this invention to further use an  
31 agent identified as described herein in an appropriate animal model. For example, an  
32 agent identified as described herein (e.g., an antigen-modulating agent, an antisense  
33 antigen nucleic acid molecule, or an antigen-specific antibody, can be used in an animal  
34 or other model to determine the efficacy, toxicity, or side effects of treatment with such an  
35 agent. Alternatively, an agent identified as described herein can be used in an animal or  
36 other model to determine the mechanism of action of such an agent. Furthermore, this  
37 invention pertains to uses of novel agents identified by the above-described screening  
38 assays for treatments as described herein.

39  
40 The antigen proteins of the present invention are also useful to provide a target for  
41 diagnosing a disease or a disease mediated by the peptide. Accordingly, the invention  
42 provides methods for detecting the presence, or levels of, the protein (or encoding  
43 mRNA) in a cell, tissue, or organism. Experimental data as provided in FIG. 22 indicates  
44 expression in T. parva-infected bovine lymphocytes. The method involves contacting a  
45 biological sample with a compound capable of interacting with the antigen protein such  
46 that the interaction can be detected. Such an assay can be provided in a single detection

1 format or a multi-detection format such as an antibody chip array.

2

3 One agent for detecting a protein in a sample is an antibody capable of selectively binding  
4 to protein. A biological sample includes tissues, cells and biological fluids isolated from a  
5 subject, as well as tissues, cells and fluids present within a subject.

6

7 The peptides of the present invention also provide targets for diagnosing active parasite  
8 activity or disease, in an animal having a variant peptide, particularly activities and  
9 conditions that are known for other members of the family of proteins to which the present  
10 one belongs. Thus, the peptide can be isolated from a biological sample and assayed for  
11 the presence of a genetic mutation that results in aberrant peptide. This includes amino  
12 acid substitution, deletion, insertion, rearrangement, (as the result of aberrant splicing  
13 events), and inappropriate post-translational modification. Analytic methods include  
14 altered electrophoretic mobility, altered tryptic peptide digest, altered enzyme activity in  
15 cell-based or cell-free assay, alteration in substrate or antibody-binding pattern, altered  
16 isoelectric point, direct amino acid sequencing, and any other of the known assay  
17 techniques useful for detecting mutations in a protein. Such an assay can be provided in  
18 a single detection format or a multi-detection format such as an antibody chip array.

19

20 In vitro techniques for detection of peptide include enzyme linked immunosorbent assays  
21 (ELISAs), Western blots, immunoprecipitations and immunofluorescence using a  
22 detection reagent, such as an antibody or protein binding agent, as well as the methods  
23 that represent embodiments of the present invention. Alternatively, the peptide can be  
24 detected in vivo in a subject by introducing into the subject a labeled anti-peptide antibody  
25 or other types of detection agent. For example, the antibody can be labeled with a  
26 radioactive marker whose presence and location in a subject can be detected by standard  
27 imaging techniques. Such methods can be used to detect the allelic variant of a peptide  
28 expressed in an infected subject and methods which detect fragments of a peptide in a  
29 sample.

30

31 The diagnostic uses can be applied, not only in genetic testing, but also in monitoring a  
32 treatment modality. Accordingly, where treatment is ultimately aimed at preventing or  
33 controlling infection, antibodies directed against the protein or relevant fragments can be  
34 used to monitor therapeutic efficacy.

35

36 The antibodies are also useful for inhibiting protein function, for example, blocking the  
37 binding of the antigenic peptide to a binding partner such as a substrate. These uses can  
38 also be applied in a therapeutic context in which treatment involves inhibiting the protein's  
39 function. An antibody can be used, for example, to block binding, thus modulating  
40 (agonizing or antagonizing) the peptides activity. Antibodies can be prepared against  
41 specific fragments containing sites required for function or against intact protein that is  
42 associated with a cell or cell membrane. See FIGS. 16-20 and 23 for structural  
43 information relating to the proteins of the present invention.

44

45 The invention also encompasses kits for using antibodies to detect the presence of a  
46 protein in a biological sample. The kit can comprise antibodies such as a labeled or

1 labelable antibody and a compound or agent for detecting protein in a biological sample;  
2 means for determining the amount of protein in the sample; means for comparing the  
3 amount of protein in the sample with a standard; and instructions for use. Such a kit can  
4 be supplied to detect a single protein or epitope or can be configured to detect one of a  
5 multitude of epitopes, such as in an antibody detection array. Arrays are described in  
6 detail below for nucleic acid arrays and similar methods have been developed for  
7 antibody arrays.

8

## 9 Nucleic Acid Molecules

10

11 [0104] The present invention further provides isolated nucleic acid molecules that encode  
12 a T. parva antigen peptide or protein of the present invention (cDNA, transcript and  
13 genomic sequence). Such nucleic acid molecules will consist of, consist essentially of, or  
14 comprise a nucleotide sequence that encodes one of the enzyme peptides of the present  
15 invention, an allelic variant thereof, or an ortholog or paralog thereof.

16

17 As used herein, an "isolated" nucleic acid molecule is one that is separated from other  
18 nucleic acid present in the natural source of the nucleic acid. Preferably, an "isolated"  
19 nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences  
20 located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from  
21 which the nucleic acid is derived. However, there can be some flanking nucleotide  
22 sequences, for example up to about 5KB, 4KB, 3KB, 2KB, or 1KB or less, particularly  
23 contiguous peptide encoding sequences and peptide encoding sequences within the  
24 same gene but separated by introns in the genomic sequence. The important point is that  
25 the nucleic acid is isolated from remote and unimportant flanking sequences such that it  
26 can be subjected to the specific manipulations described herein such as recombinant  
27 expression, preparation of probes and primers, and other uses specific to the nucleic acid  
28 sequences.

29

30 Moreover, an "isolated" nucleic acid molecule, such as a transcript/cDNA molecule, can  
31 be substantially free of other cellular material, or culture medium when produced by  
32 recombinant techniques, or chemical precursors or other chemicals when chemically  
33 synthesized. However, the nucleic acid molecule can be fused to other coding or  
34 regulatory sequences and still be considered isolated.

35

36 For example, recombinant DNA molecules contained in a vector are considered isolated.  
37 Further examples of isolated DNA molecules include recombinant DNA molecules  
38 maintained in heterologous host cells or purified (partially or substantially) DNA molecules  
39 in solution. Isolated RNA molecules include in vivo or in vitro RNA transcripts of the  
40 isolated DNA molecules of the present invention. Isolated nucleic acid molecules  
41 according to the present invention further include such molecules produced synthetically.

42

43 Accordingly, the present invention provides nucleic acid molecules that consist of the  
44 nucleotide sequence shown in FIGS. 16-20 or any nucleic acid molecule that encodes the  
45 protein (amino acid sequence) included in these figures. A nucleic acid molecule consists  
46 of a nucleotide sequence when the nucleotide sequence is the complete nucleotide

1 sequence of the nucleic acid molecule.

2

3 The present invention further provides nucleic acid molecules that consist essentially of  
4 the nucleotide sequence shown in FIGS. 16-20, or any nucleic acid molecule that  
5 encodes the proteins provided in the same figures. A nucleic acid molecule consists  
6 essentially of a nucleotide sequence when such a nucleotide sequence is present with  
7 only a few additional nucleic acid residues in the final nucleic acid molecule.

8

9 The present invention further provides nucleic acid molecules that comprise the  
10 nucleotide sequences shown shown in FIGS. 16-20, or any nucleic acid molecule that  
11 encodes the proteins provided in the same figures A nucleic acid molecule comprises a  
12 nucleotide sequence when the nucleotide sequence is at least part of the final nucleotide  
13 sequence of the nucleic acid molecule. In such a fashion, the nucleic acid molecule can  
14 be only the nucleotide sequence or have additional nucleic acid residues, such as nucleic  
15 acid residues that are naturally associated with it or heterologous nucleotide sequences.  
16 Such a nucleic acid molecule can have a few additional nucleotides or can comprises  
17 several hundred or more additional nucleotides. A brief description of how various types  
18 of these nucleic acid molecules can be readily made/isolated is provided below.

19

20

21 The invention further provides nucleic acid molecules that encode fragments of the  
22 peptides of the present invention as well as nucleic acid molecules that encode obvious  
23 variants of the enzyme proteins of the present invention that are described above. Such  
24 nucleic acid molecules may be naturally occurring, such as allelic variants (same locus),  
25 paralogs (different locus), and orthologs (different organism), or may be constructed by  
26 recombinant DNA methods or by chemical synthesis. Such non-naturally occurring  
27 variants may be made by mutagenesis techniques, including those applied to nucleic acid  
28 molecules, cells, or organisms. Accordingly, as discussed above, the variants can contain  
29 nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either  
30 or both the coding and non-coding regions. The variations can produce both conservative  
31 and non- conservative amino acid substitutions.

32

33

34 The present invention further provides epitope fragments of the antigen proteins and the  
35 nucleic acid molecules encoding the antigens, provided in the description of epitope-  
36 mapping experiments described below. An epitope coding region comprises a contiguous  
37 nucleotide sequence greater than 12 or more nucleotides. Further, a fragment could at  
38 least 30, 40, 50, 100, 250 or 500 nucleotides in length. The length of the fragment will be  
39 based on its intended use. For example, the fragment can encode epitope bearing  
40 regions of the peptide, or can be useful as DNA probes and primers. Such fragments can  
41 be isolated using the known nucleotide sequence to synthesize an oligonucleotide probe.  
42 A labeled probe can then be used to screen a cDNA library, genomic DNA library, or  
43 mRNA to isolate nucleic acid corresponding to the coding region. Further, primers can be  
44 used in PCR reactions to clone specific regions of gene.

45

46



1 A probe/primer typically comprises substantially a purified oligonucleotide or  
2 oligonucleotide pair. The oligonucleotide typically comprises a region of nucleotide  
3 sequence that hybridizes under stringent conditions to at least about 12, 20, 25, 40, 50 or  
4 more consecutive nucleotides.

5  
6  
7  
8  
9  
10  
11 Orthologs, homologs, and allelic variants can be identified using methods well known in  
12 the art. As described in the Peptide Section, these variants comprise a nucleotide  
13 sequence encoding a peptide that is typically 60-70%, 70-80%, 80-90%, and more  
14 typically at least about 90-95% or more homologous to the nucleotide sequence shown in  
15 the Figure sheets or a fragment of this sequence. Such nucleic acid molecules can readily  
16 be identified as being able to hybridize under moderate to stringent conditions, to the  
17 nucleotide sequence shown in the Figure sheets or a fragment of the sequence. Allelic  
18 variants can readily be determined by genetic locus of the encoding gene. The gene  
19 encoding the novel enzyme of the present invention is located on a genome component  
20 that has been mapped to the *T. parva* chromosome1 and 2 which is supported by multiple  
21 lines of evidence, such as STS and BAC map data.

#### 22 23 24 Nucleic Acid Molecule Uses

25  
26 The nucleic acid molecules of the present invention are useful for probes, primers,  
27 chemical intermediates, and in biological assays. The nucleic acid molecules are useful  
28 as a hybridization probe for messenger RNA, transcript/cDNA and genomic DNA to  
29 isolate full-length cDNA and genomic clones encoding the peptide described in FIG. 16-  
30 20, and to isolate cDNA and genomic clones that correspond to variants (alleles,  
31 orthologs, etc.) producing the same or related peptides shown in FIG. 16-20 and FIG 23.  
32 As illustrated in FIG. 11, deletions were identified at 2 different nucleotide positions.

33  
34 The probe can correspond to any sequence along the entire length of the nucleic acid  
35 molecules provided in the Figures. Accordingly, it could be derived from 5' noncoding  
36 regions, the coding region, and 3' noncoding regions. However, as discussed, fragments  
37 are not to be construed as encompassing fragments disclosed prior to the present  
38 invention.

39  
40 The nucleic acid molecules are also useful as primers for PCR to amplify any given region  
41 of a nucleic acid molecule and are useful to synthesize antisense molecules of desired  
42 length and sequence.

43  
44 The nucleic acid molecules are also useful for constructing recombinant vectors. Such  
45 vectors include expression vectors that express a portion of, or all of, the peptide  
46 sequences. Vectors also include insertion vectors, used to integrate into another nucleic

1 acid molecule sequence, such as into the cellular genome, to alter in situ expression of a  
2 gene and/or gene product. For example, an endogenous coding sequence can be  
3 replaced via homologous recombination with all or part of the coding region containing  
4 one or more specifically introduced mutations.

5  
6  
7 The nucleic acid molecules are also useful for expressing antigenic portions of the  
8 proteins.

9  
10 The nucleic acid molecules are also useful as probes for determining the chromosomal  
11 positions of the nucleic acid molecules by means of in situ hybridization methods. The  
12 genes encoding the novel antigens of the present invention are located on a genome  
13 component that has been mapped to *T. parva* chromosomes 1 (Tp1 & Tp4) and 2 (Tp5,  
14 Tp7 & Tp8) .

15  
16 The nucleic acid molecules are also useful in making vectors containing the gene  
17 regulatory regions of the nucleic acid molecules of the present invention.

18  
19 The nucleic acid molecules are also useful for designing ribozymes/antigens  
20 corresponding to all, or a part, of the mRNA produced from the nucleic acid molecules  
21 described herein.

22  
23 The nucleic acid molecules are also useful for making vectors that express part, or all, of  
24 the peptides.

25  
26 The nucleic acid molecules are also useful for constructing host cells expressing a part, or  
27 all, of the nucleic acid molecules and peptides.

28  
29 The nucleic acid molecules are also useful for constructing transgenic animals expressing  
30 all, or a part, of the nucleic acid molecules and peptides.

31  
32 The nucleic acid molecules are also useful as hybridization probes for determining the  
33 presence, level, form and distribution of nucleic acid expression. Such probes could be  
34 used to detect the presence of, or to determine levels of, a specific nucleic acid molecule  
35 in cells, tissues, and in organisms. The nucleic acid whose level is determined can be  
36 DNA or RNA. Accordingly, probes corresponding to the peptides described herein can be  
37 used to assess expression and/or gene copy number in a given cell, tissue, or organism.  
38 These uses are relevant for diagnosis of disorders involving an increase or decrease in  
39 antigen protein expression relative to normal results.

40  
41 In vitro techniques for detection of mRNA include Northern hybridizations and in situ  
42 hybridizations. In vitro techniques for detecting DNA includes Southern hybridizations and  
43 in situ hybridization.

44  
45 Probes can be used as a part of a diagnostic test kit for identifying cells or tissues that  
46 express an *T. parva* antigen protein, such as by measuring a level of an antigen-encoding

1 nucleic acid in a sample of cells from a subject e.g., mRNA or genomic DNA, or  
2 determining if an antigen gene has been mutated.  
3  
4 Nucleic acid expression assays are useful for drug screening to identify compounds that  
5 modulate protozoan antigen nucleic acid expression.  
6  
7 The invention thus provides a method for identifying a compound that can be used to treat  
8 a disease associated with nucleic acid expression of the antigen gene, particularly  
9 biological and pathological processes that are mediated by the presence of *T. parva* in  
10 cells and tissues that express it. The method typically includes assaying the ability of the  
11 compound to modulate the expression of the antigen nucleic acid and thus identifying a  
12 compound that can be used to treat a disorder characterized by nucleic acid expression  
13 by an infecting protozoan. The assays can be performed in cell-based and cell-free  
14 systems. Cell-based assays include cells naturally expressing the enzyme nucleic acid or  
15 recombinant cells genetically engineered to express specific nucleic acid sequences.  
16  
17 The assay for enzyme nucleic acid expression can involve direct assay of nucleic acid  
18 levels, such as mRNA levels, or on collateral compounds involved in an antigen-specific  
19 signal pathway. Further, the expression of genes that are up- or down-regulated in  
20 response to the antigen-stimulating signal pathway can also be assayed. In this  
21 embodiment the regulatory regions of these genes can be operably linked to a reporter  
22 gene such as luciferase.  
23  
24 Thus, modulators of antigen gene expression can be identified in a method wherein a cell  
25 is contacted with a candidate compound and the expression of mRNA determined. The  
26 level of expression of antigen mRNA in the presence of the candidate compound is  
27 compared to the level of expression of antigen mRNA in the absence of the candidate  
28 compound. The candidate compound can then be identified as a modulator of nucleic  
29 acid expression based on this comparison and be used, for example to treat an infection  
30 by a pathogenic protozoan. When expression of mRNA is statistically significantly greater  
31 in the presence of the candidate compound than in its absence, the candidate compound  
32 is identified as a stimulator of nucleic acid expression. When nucleic acid expression is  
33 statistically significantly less in the presence of the candidate compound than in its  
34 absence, the candidate compound is identified as an inhibitor of nucleic acid expression.  
35  
36  
37 The nucleic acid molecules are also useful for monitoring the effectiveness of modulating  
38 compounds on the infectivity of a protozoan in clinical trials or in a treatment regimen.  
39 Thus, the gene expression pattern can serve as a barometer for the continuing  
40 effectiveness of treatment with the compound, particularly with compounds to which an  
41 animal or the pathogen can develop resistance. The gene expression pattern can also  
42 serve as a marker indicative of a physiological response of the affected cells to the  
43 compound.  
44  
45 Accordingly, such monitoring would allow either increased administration of the  
46 compound or the administration of alternative compounds to which the parasite has not

1 become resistant. Similarly, if the level of nucleic acid expression falls below a desirable  
2 level, administration of the compound could be commensurately decreased.

3  
4 The nucleic acid molecules are also useful in diagnostic assays for qualitative changes in  
5 antigen nucleic acid expression, and particularly in qualitative changes that lead to  
6 pathology. The nucleic acid molecules can be used to detect mutations in antigen genes  
7 and gene expression products such as mRNA. The nucleic acid molecules can be used  
8 as hybridization probes to detect naturally occurring genetic mutations in the antigen gene  
9 and thereby to determine whether a parasite with the mutation is of a different strain or  
10 has a differential pathogenic potential. Mutations include deletion, addition, or substitution  
11 of one or more nucleotides in the gene, chromosomal rearrangement, such as inversion  
12 or transposition, modification of genomic DNA, such as aberrant methylation patterns or  
13 changes in gene copy number, such as amplification. Detection of a mutated form of the  
14 antigen gene associated with a dysfunction provides a diagnostic tool for an active  
15 disease or susceptibility to disease when the disease results from over-expression,  
16 under-expression, or altered expression of an antigen protein.

17  
18 Parasite strains carrying mutations in the enzyme gene can be detected at the nucleic  
19 acid level by a variety of techniques. SNPs that have been found in the gene encoding  
20 the antigen of the present invention. SNPs were identified at 4 different nucleotide  
21 positions. Genomic DNA can be analyzed directly or can be amplified by using PCR prior  
22 to analysis. RNA or cDNA can be used in the same way. In some uses, detection of the  
23 mutation involves the use of a probe/primer in a polymerase chain reaction (PCR) (see,  
24 e.g. U.S. Pat. Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or,  
25 alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al., Science  
26 241:1077-1080 (1988); and Nakazawa et al., PNAS 91:360-364 (1994)), the latter of  
27 which can be particularly useful for detecting point mutations in the gene (see Abravaya  
28 et al., Nucleic Acids Res. 23:675-682 (1995)). This method can include the steps of  
29 collecting a sample of cells from an infective parasite, isolating nucleic acid (e.g.,  
30 genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample  
31 with one or more primers which specifically hybridize to a gene under conditions such that  
32 hybridization and amplification of the gene (if present) occurs, and detecting the presence  
33 or absence of an amplification product, or detecting the size of the amplification product  
34 and comparing the length to a control sample. Deletions and insertions can be detected  
35 by a change in size of the amplified product compared to the normal genotype. Point  
36 mutations can be identified by hybridizing amplified DNA to normal RNA or antisense  
37 DNA sequences.

38  
39 Alternatively, mutations in a enzyme gene can be directly identified, for example, by  
40 alterations in restriction enzyme digestion patterns determined by gel electrophoresis.

41  
42 Further, sequence-specific ribozymes (U.S. Pat. No. 5,498,531) can be used to score for  
43 the presence of specific mutations by development or loss of a ribozyme cleavage site.  
44 Perfectly matched sequences can be distinguished from mismatched sequences by  
45 nuclease cleavage digestion assays or by differences in melting temperature.

1 Sequence changes at specific locations can also be assessed by nuclease protection  
2 assays such as RNase and S1 protection or the chemical cleavage method. Furthermore,  
3 sequence differences between a mutant enzyme gene and a wild-type gene can be  
4 determined by direct DNA sequencing. A variety of automated sequencing procedures  
5 can be utilized when performing the diagnostic assays (Naeve, C. W., (1995)  
6 *Biotechniques* 19:448), including sequencing by mass spectrometry (see, e.g., PCT  
7 International Publication No. WO94/16101; Cohen et al., *Adv. Chromatogr.* 36:127-162  
8 (1996); and Griffin et al., *Appl. Biochem. Biotechnol.* 38:147-159 (1993)).

9  
10 Other methods for detecting mutations in the gene include methods in which protection  
11 from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA  
12 duplexes (Myers et al., *Science* 230:1242 (1985)); Cotton et al., *PNAS* 85:4397 (1988);  
13 Saleeba et al., *Meth. Enzymol.* 217:286-295 (1992)), electrophoretic mobility of mutant  
14 and wild type nucleic acid is compared (Orita et al., *PNAS* 86:2766 (1989); Cotton et al.,  
15 *Mutat. Res.* 285:125-144 (1993); and Hayashi et al., *Genet. Anal. Tech. Appl.* 9:73-79  
16 (1992)), and movement of mutant or wild-type fragments in polyacrylamide gels  
17 containing a gradient of denaturant is assayed using denaturing gradient gel  
18 electrophoresis (Myers et al., *Nature* 313:495 (1985)). Examples of other techniques for  
19 detecting point mutations include selective oligonucleotide hybridization, selective  
20 amplification, and selective primer extension.

21  
22 The nucleic acid molecules are also useful for testing an individual parasite for a  
23 genotype that while not necessarily causing the disease, nevertheless affects the  
24 treatment modality. Thus, the nucleic acid molecules can be used to study the  
25 relationship between a particular strain of parasite's genotype and the parasite's response  
26 to a compound used for treatment (pharmacogenomic relationship). Accordingly, the  
27 nucleic acid molecules described herein can be used to assess the mutation content of  
28 the antigen gene in an individual in order to select an appropriate compound or dosage  
29 regimen for treatment. FIG. 23 provides information on SNPs that have been found in  
30 eleven strains in the gene encoding the antigen of the present invention. SNPs were  
31 identified at nucleotide positions underlined in FIG 23. Some of these SNPs that are  
32 located outside the ORF and in introns may affect gene transcription.

33  
34 Thus nucleic acid molecules displaying genetic variations that affect treatment provide a  
35 diagnostic target that can be used to tailor treatment in an individual. Accordingly, the  
36 production of recombinant cells and animals containing these polymorphisms allow  
37 effective clinical design of treatment compounds and dosage regimens.

38  
39 The nucleic acid molecules are thus useful as antisense constructs to control antigen  
40 gene expression in cells, tissues, and organisms. A DNA antisense nucleic acid molecule  
41 is designed to be complementary to a region of the gene involved in transcription,  
42 preventing transcription and hence production of enzyme protein. An antisense RNA or  
43 DNA nucleic acid molecule would hybridize to the mRNA and thus block translation of  
44 mRNA into antigen protein. Alternatively, a class of antisense molecules can be used to  
45 inactivate mRNA in order to decrease expression of antigen nucleic acid.

1  
2  
3 The invention also encompasses kits for detecting the presence of an antigen nucleic acid  
4 in a biological sample. For example, the kit can comprise reagents such as a labeled or  
5 labelable nucleic acid or agent capable of detecting antigen nucleic acid in a biological  
6 sample; means for determining the amount of antigen nucleic acid in the sample; and  
7 means for comparing the amount of antigen nucleic acid in the sample with a standard.  
8 The compound or agent can be packaged in a suitable container. The kit can further  
9 comprise instructions for using the kit to detect antigen protein mRNA or DNA.

#### 10 11 Nucleic Acid Arrays

12  
13 The present invention further provides nucleic acid detection kits, such as arrays or  
14 microarrays of nucleic acid molecules that are based on the sequence information  
15 provided in FIGS. 16-20.

16  
17 As used herein "Arrays" or "Microarrays" refers to an array of distinct polynucleotides or  
18 oligonucleotides synthesized on a substrate, such as paper, nylon or other type of  
19 membrane, filter, chip, glass slide, or any other suitable solid support. In one  
20 embodiment, the microarray is prepared and used according to the methods described in  
21 U.S. Pat. No. 5,837,832, Chee et al., PCT application W095/11995 (Chee et al.),  
22 Lockhart, D. J. et al. (1996; Nat. Biotech. 14: 1675-1680) and Schena, M. et al. (1996;  
23 Proc. Natl. Acad. Sci. 93: 10614-10619), all of which are incorporated herein in their  
24 entirety by reference. In other embodiments, such arrays are produced by the methods  
25 described by Brown et al., U.S. Pat. No. 5,807,522.

26  
27 The microarray or detection kit is preferably composed of a large number of unique,  
28 single-stranded nucleic acid sequences, usually either synthetic antisense  
29 oligonucleotides or fragments of cDNAs, fixed to a solid support. The oligonucleotides are  
30 preferably about 6-60 nucleotides in length, more preferably 15-30 nucleotides in length,  
31 and most preferably about 20-25 nucleotides in length. For a certain type of microarray or  
32 detection kit, it may be preferable to use oligonucleotides that are only 7-20 nucleotides in  
33 length. The microarray or detection kit may contain oligonucleotides that cover the known  
34 5', or 3', sequence, sequential oligonucleotides which cover the full length sequence; or  
35 unique oligonucleotides selected from particular areas along the length of the sequence.  
36 Polynucleotides used in the microarray or detection kit may be oligonucleotides that are  
37 specific to a gene or genes of interest.

38  
39 In order to produce oligonucleotides to a known sequence for a microarray or detection  
40 kit, the gene(s) of interest (or an ORF identified in the present invention) is typically  
41 examined using a computer algorithm which starts at the 5' or at the 3' end of the  
42 nucleotide sequence. Typical algorithms will then identify oligomers of defined length that  
43 are unique to the gene, have a GC content within a range suitable for hybridization, and  
44 lack predicted secondary structure that may interfere with hybridization. In certain  
45 situations it may be appropriate to use pairs of oligonucleotides on a microarray or  
46 detection kit. The "pairs" will be identical, except for one nucleotide that preferably is

1 located in the center of the sequence. The second oligonucleotide in the pair  
2 (mismatched by one) serves as a control. The number of oligonucleotide pairs may range  
3 from two to one million. The oligomers are synthesized at designated areas on a  
4 substrate using a light-directed chemical process. The substrate may be paper, nylon or  
5 other type of membrane, filter, chip, glass slide or any other suitable solid support.

6  
7 In another aspect, an oligonucleotide may be synthesized on the surface of the substrate  
8 by using a chemical coupling procedure and an ink jet application apparatus, as  
9 described in PCT application W095/251116 (Baldeschweiler et al.) which is incorporated  
10 herein in its entirety by reference. In another aspect, a "gridded" array analogous to a dot  
11 (or slot) blot may be used to arrange and link cDNA fragments or oligonucleotides to the  
12 surface of a substrate using a vacuum system, thermal, UV, mechanical or chemical  
13 bonding procedures. An array, such as those described above, may be produced by hand  
14 or by using available devices (slot blot or dot blot apparatus), materials (any suitable solid  
15 support), and machines (including robotic instruments), and may contain 8, 24, 96, 384,  
16 1536, 6144 or more oligonucleotides, or any other number between two and one million  
17 which lends itself to the efficient use of commercially available instrumentation.

18  
19 In order to conduct sample analysis using a microarray or detection kit, the RNA or DNA  
20 from a biological sample is made into hybridization probes. The mRNA is isolated, and  
21 cDNA is produced and used as a template to make antisense RNA (aRNA). The aRNA is  
22 amplified in the presence of fluorescent nucleotides, and labeled probes are incubated  
23 with the microarray or detection kit so that the probe sequences hybridize to  
24 complementary oligonucleotides of the microarray or detection kit. Incubation conditions  
25 are adjusted so that hybridization occurs with precise complementary matches or with  
26 various degrees of less complementarity. After removal of nonhybridized probes, a  
27 scanner is used to determine the levels and patterns of fluorescence. The scanned  
28 images are examined to determine degree of complementarity and the relative  
29 abundance of each oligonucleotide sequence on the microarray or detection kit. The  
30 biological samples may be obtained from any bodily fluids (such as blood, urine, saliva,  
31 phlegm, gastric juices, etc.), cultured cells, biopsies, or other tissue preparations. A  
32 detection system may be used to measure the absence, presence, and amount of  
33 hybridization for all of the distinct sequences simultaneously. This data may be used for  
34 large-scale correlation studies on the sequences, expression patterns, mutations,  
35 variants, or polymorphisms among samples.

36  
37 Using such arrays, the present invention provides methods to identify the expression of  
38 the antigen proteins/peptides of the present invention. In detail, such methods comprise  
39 incubating a test sample with one or more nucleic acid molecules and assaying for  
40 binding of the nucleic acid molecule with components within the test sample. Such assays  
41 will typically involve arrays comprising many genes, at least one of which is a gene of the  
42 present invention and or alleles of the enzyme gene of the present invention. The figures  
43 and associated information below provide information on epitope sequence and micro-  
44 variation in strains of *T. parva*, that have been found in the gene encoding the antigen of  
45 the present invention.

1  
2 Conditions for incubating a nucleic acid molecule with a test sample vary. Incubation  
3 conditions depend on the format employed in the assay, the detection methods  
4 employed, and the type and nature of the nucleic acid molecule used in the assay. One  
5 skilled in the art will recognize that any one of the commonly available hybridization,  
6 amplification or array assay formats can readily be adapted to employ the novel  
7 fragments of the T. parva genome disclosed herein. Examples of such assays can be  
8 found in Chard, T, An Introduction to Radioimmunoassay and Related Techniques,  
9 Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G. R. et al.,  
10 Techniques in Immunocytochemistry, Academic Press, Orlando, Fla. Vol. 1 (1982), Vol. 2  
11 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of Enzyme Immunoassays:  
12 Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science  
13 Publishers, Amsterdam, The Netherlands (1985).

14  
15 The test samples of the present invention include cells, protein or membrane extracts of  
16 cells. The test sample used in the above-described method will vary based on the assay  
17 format, nature of the detection method and the tissues, cells or extracts used as the  
18 sample to be assayed. Methods for preparing nucleic acid extracts or of cells are well  
19 known in the art and can be readily be adapted in order to obtain a sample that is  
20 compatible with the system utilized.

21  
22 In another embodiment of the present invention, kits are provided which contain the  
23 necessary reagents to carry out the assays of the present invention.

24  
25 Specifically, the invention provides a compartmentalized kit to receive, in close  
26 confinement, one or more containers which comprises: (a) a first container comprising  
27 one of the nucleic acid molecules that can bind to a fragment of the T. parva genome  
28 disclosed herein; and (b) one or more other containers comprising one or more of the  
29 following: wash reagents, reagents capable of detecting presence of a bound nucleic  
30 acid.

31  
32 In detail, a compartmentalized kit includes any kit in which reagents are contained in  
33 separate containers. Such containers include small glass containers, plastic containers,  
34 strips of plastic, glass or paper, or arraying material such as silica. Such containers allows  
35 one to efficiently transfer reagents from one compartment to another compartment such  
36 that the samples and reagents are not cross-contaminated, and the agents or solutions of  
37 each container can be added in a quantitative fashion from one compartment to another.  
38 Such containers will include a container which will accept the test sample, a container  
39 which contains the nucleic acid probe, containers which contain wash reagents (such as  
40 phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents  
41 used to detect the bound probe. One skilled in the art will readily recognize that the  
42 previously unidentified enzyme gene of the present invention can be routinely identified  
43 using the sequence information disclosed herein can be readily incorporated into one of  
44 the established kit formats which are well known in the art, particularly expression arrays.



1  
2 Genetically engineered host cells can be further used to produce non-human transgenic  
3 animals. A transgenic animal is preferably a mammal, for example a rodent, such as a rat  
4 or mouse, in which one or more of the cells of the animal include a transgene. A  
5 transgene is exogenous DNA which is integrated into the genome of a cell from which a  
6 transgenic animal develops and which remains in the genome of the mature animal in one  
7 or more cell types or tissues of the transgenic animal. These animals are useful for  
8 studying the function of an antigen protein and identifying and evaluating modulators of  
9 parasite development or activity. Other examples of transgenic animals include non-  
10 human primates, sheep, dogs, cows, goats, chickens, and amphibians.

11  
12 A transgenic animal can be produced by introducing nucleic acid into the male pronuclei  
13 of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to  
14 develop in a pseudopregnant female foster animal. Any of the enzyme protein nucleotide  
15 sequences can be introduced as a transgene into the genome of a non-human animal,  
16 such as a mouse.

17  
18 Any of the regulatory or other sequences useful in expression vectors can form part of the  
19 transgenic sequence. This includes intronic sequences and polyadenylation signals, if not  
20 already included. A tissue-specific regulatory sequence(s) can be operably linked to the  
21 transgene to direct expression of the enzyme protein to particular cells.

22  
23 Methods for generating transgenic animals via embryo manipulation and microinjection,  
24 particularly animals such as mice, have become conventional in the art and are  
25 described, for example, in U.S. Pat. Nos. 4,736,866 and 4,870,009, both by Leder et al.,  
26 U.S. Pat. No. 4,873,191 by Wagner et al. and in Hogan, B., Manipulating the Mouse  
27 Embryo, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar  
28 methods are used for production of other transgenic animals. A transgenic founder animal  
29 can be identified based upon the presence of the transgene in its genome and/or  
30 expression of transgenic mRNA in tissues or cells of the animals. A transgenic founder  
31 animal can then be used to breed additional animals carrying the transgene. Moreover,  
32 transgenic animals carrying a transgene can further be bred to other transgenic animals  
33 carrying other transgenes. A transgenic animal also includes animals in which the entire  
34 animal or tissues in the animal have been produced using the homologously recombinant  
35 host cells described herein.

36  
37 In another embodiment, transgenic non-human animals can be produced which contain  
38 selected systems that allow for regulated expression of the transgene. One example of  
39 such a system is the cre/loxP recombinase system of bacteriophage P1. For a description  
40 of the cre/loxP recombinase system, see, e.g., Lakso et al. PNAS 89:6232-6236 (1992).  
41 Another example of a recombinase system is the FLP recombinase system of *S.*  
42 *cerevisiae* (O'Gorman et al. Science 251:1351-1355 (1991)). If a cre/loxP recombinase  
43 system is used to regulate expression of the transgene, animals containing transgenes  
44 encoding both the Cre recombinase and a selected protein is required. Such animals can  
45 be provided through the construction of "double" transgenic animals, e.g., by mating two  
46 transgenic animals, one containing a transgene encoding a selected protein and the other

1 containing a transgene encoding a recombinase.

2

3 Clones of the non-human transgenic animals described herein can also be produced  
4 according to the methods described in Wilmut, I. et al. Nature 385:810-813 (1997) and  
5 PCT International Publication Nos. WO97/07668 and WO97/07669. In brief, a cell, e.g., a  
6 somatic cell, from the transgenic animal can be isolated and induced to exit the growth,  
7 cycle and enter G.sub.o phase. The quiescent cell can then be fused, e.g., through the  
8 use of electrical pulses, to an enucleated oocyte from an animal of the same species from  
9 which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it  
10 develops to morula or blastocyst and then transferred to pseudopregnant female foster  
11 animal. The offspring born of this female foster animal will be a clone of the animal from  
12 which the cell, e.g., the somatic cell, is isolated.

13

14 Transgenic animals containing recombinant cells that express the peptides described  
15 herein are useful to conduct the assays described herein in an in vivo context.  
16 Accordingly, the various physiological factors that are present in vivo and that could effect  
17 substrate binding, enzyme protein activation, and signal transduction, may not be evident  
18 from in vitro cell-free or cell-based assays. Accordingly, it is useful to provide non-human  
19 transgenic animals to assay in vivo enzyme protein function, including substrate  
20 interaction, the effect of specific mutant enzyme proteins on enzyme protein function and  
21 substrate interaction, and the effect of chimeric enzyme proteins. It is also possible to  
22 assess the effect of null mutations, that is, mutations that substantially or completely  
23 eliminate one or more enzyme protein functions.

24

25 The T. parva antigens of the present invention have particular potential for induction of in  
26 vivo, antigen-specific CD8+ cytotoxic T cells for prophylactic immunization of cattle for the  
27 prevention of East Coast Fever disease. In addition, CTLs specific to the following  
28 metazoan parasites may also be induced by compositions identified in accordance with  
29 the methods of the present invention: Plasmodium falciparum (which causes malaria),  
30 Schistosoma mansoni (which causes schistosomiasis), and Trypanosoma cruzi (which  
31 causes Chagas' disease), Giardia lamblia, Entamoeba histolytica, Cryptosporidium spp.,  
32 Leishmania spp., Brugia spp., Wuchereria spp., Onchocerca spp., Strongyloides spp.,  
33 Coccidia, Haemaphysalis spp., Ostertagia spp., Trichomonas spp., Dirofilaria spp.,  
34 Toxocara spp., Naegleria spp., Pneumocystis carinii, Ascaris spp., other Trypanosoma  
35 spp., other Schistosome spp., other Plasmodium spp., Babesia spp., Theileria spp.,  
36 including but not limited to T. parva, T. lawrencei, T. annulata, T. hirci, T. ovis, T.  
37 lastoguardi, T. orientalis, T. buffeli and T. taurotragi, Babesia spp., including, but not  
38 limited to B. bigemina, B. divergens, B. major, B. bovis, B. motasi, B. ovis, B. cabelli, B.  
39 equi, B. traumani, B. canis, B. gibsoni, B. felis, and B. microfti, Adelina spp., including,  
40 but not limited to A. delina, A. castana, A. picei, A. palori, and A. triboli, Anisakis and  
41 Isospora belli.

42

**Vaccination strategy**

Cattle trials have been initiated to assess the vaccine potential of identified CTL target antigens utilizing a recombinant canarypox virus (patented by Merial Ltd) as an antigen delivery method. The first of these trials is testing one of the candidate antigens and involves use of 16 cattle. Animals were inoculated intramuscularly with 1 ml of vaccine ( $1 \times 10^8$  pfu of virus) and boosted similarly after 4 weeks. Following a further 4 weeks, cattle will be subjected to an LD<sub>100</sub> challenge with *T.parva* sporozoites by administering subcutaneously 1 ml of diluted stabilated infective material. Animals will be monitored parasitologically and clinically over a period of 2-3 weeks to determine whether the vaccine has protected.

Immunological assays will be performed following immunization and challenge to evaluate antigen-specific CTL responses and relate these to the outcome to challenge.

All publications and patents mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described method and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the above-described modes for carrying out the invention which are obvious to those skilled in the field of molecular biology or related fields are intended to be within the scope of the following outline of the invention ("claims").

**The invention is further described in the attached Annexes:**

**ANNEX A: CTL TARGET ANTIGEN IDENTIFICATION METHOD**

**ANNEX B: IDENTIFICATION OF A VACCINE CANDIDATE ANTIGEN RECOGNISED BY *THEILERIA PARVA* SPECIFIC BOVINE CYTOTOXIC T LYMPHOCYTES**

## 1 OUTLINE OF THE INVENTION:

- 2 1) A method for the identification of parasite antigens that are targets of cytotoxic T  
3 lymphocytes, comprising the step of: co-culturing immortalized fibroblast cell lines  
4 transfected with pooled cDNA harvested from a pathogen, with clones of lines of cytotoxic T  
5 cells, generated in an animal that has been immunized, by infection and treatment with the  
6 pathogen.
- 7 2) A method for a three-way matrix resolution for identification of a single cDNA clone from  
8 a pool of cDNAs, in high throughput procedures, comprising the steps of:
- 9 a) transforming bacterial cells with DNA from a pool of about 25 to about 500  
10 cDNAs, that has tested positive in a routine assay;
  - 11 b) diluting the culture of transformed cells so as to yield a density of about 500-5000  
12 colonies per 150 cm<sup>2</sup>, when plated on agar-containing plates;
  - 13 c) picking about 100 to 500 colonies from growth cultures;
  - 14 d) placing about 5 to 60 pools of about 10-100 individual cultures grown from the  
15 colonies, into numbered tubes, in such a manner such that each individual bacterial  
16 culture is present in 3 of said pools, so that tubes are labelled with a number and  
17 positioned so that matrices of tubes are created so as to accommodate a multi-  
18 channel pipetting device;
  - 19 e) creating a corresponding matrix table is by arraying the numbers on the  
20 corresponding tubes containing the pools into an matrix table;
  - 21 f) testing the DNA from each of the tubes in a screening assay; and,
- 22 identifying the individual positive colony by comparing the results with the array.
- 23 2) The method of claim 2, wherein the screening assay causes the release of gamma interferon  
24 by CD8<sup>+</sup> cytotoxic T cells.
- 25 3) A combination of Claim 1 and Claim 2, further comprising the step of: assaying the  
26 supernatant of co-cultured cells, for the presence of soluble factors, secreted by the cytotoxic  
27 T cells.
- 28 4) The method as claimed in claim 1, wherein the pathogen is a protozoan organism.
- 29 6) The method as claimed in claim 2, wherein the protozoan organism is the macroscizhont  
30 stage of an organism in the genus Theileria.
- 31 3) The method as claimed in claim, 6 wherein the organism is Theileria parva.

- 1 4) The method as claimed in claim 4, wherein the soluble factor is a cytokine.
- 2 5) The method as claimed in claim 8, wherein the cytokine is gamma interferon.
- 3 6) An isolated polynucleotide represented by SEQ ID NO:1. (Information in Fig. 16)
- 4 7) An isolated polynucleotide represented by SEQ ID NO:2. (Information in Fig. 17)
- 5 8) An isolated polynucleotide represented by SEQ ID NO:3. (Information in Fig. 18)
- 6 9) An isolated polynucleotide represented by SEQ ID NO:4. (Information in Fig. 19)
- 7 10) An isolated polynucleotide represented by SEQ ID NO:5. (Information in Fig. 20)
- 8 11) An isolated polypeptide represented by SEQ ID NO: 7. (Information in FIG 16.)
- 9 12) An isolated polypeptide represented by SEQ ID NO: 8. (Information in FIG 17.)
- 10 13) An isolated polypeptide represented by SEQ ID NO: 9. (Information in FIG 18.)
- 11 14) An isolated polypeptide represented by SEQ ID NO: 10. (Information in FIG 19.)
- 12 15) An isolated polypeptide represented by SEQ ID NO: 11 (Information in FIG 20)
- 13 16) An isolated polypeptide represented by SEQ ID NO:12 (Information in FIG. 10, as Tp1
- 14 Del1, wherein said sequence stimulates cytotoxic T cells.
- 15 17) An isolated polypeptide represented by SEQ ID NO: 13, (Information in FIG. 11, as Tp1
- 16 Del2, wherein said sequence stimulates cytotoxic T cells.
- 17 18) An isolated isolated polypeptide of the following sequence:VGYPKVKEEML.
- 18 19) A method for generating polyclonal antibody in an animal wherein the immunizing antigen is
- 19 the polypeptide represented by the sequence of one of claim 15-21.
- 20 20) A method for generating monoclonal antibody, wherein the immunizing antigen is the
- 21 polypeptide represented by the sequence of one of that of claims 15-21.
- 22 21) A method comprising using of the polyclonal antibody or specific binding parts thereof, that
- 23 are products of claim 23 to identify the presence of a T. parva antigen in tissue harvested
- 24 from an animal.
- 25 22) The use of the polyclonal antibody or specific binding parts thereof, that are products of
- 26 claim 23 to quantify the expression of T. parva antigen in cells or biologic fluids.
- 27 23) A diagnostic kit, for the detection of protozoan infection, using the polyclonal antibody or
- 28 specific binding parts thereof, that are products of claim 23.

- 1 24) The use of the polyclonal antibody or specific binding parts thereof, that are products of  
2 claim 23 in a Western blot assay.
- 3 25) The use of the polyclonal antibody or specific binding parts thereof, that are products of  
4 claim 23 in an ELISA assay.
- 5 26) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
6 claim 24 in detection of specific molecules using a flow microfluorimeter.
- 7 27) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
8 claim 24 to purify antigen that binds to the antibody or its parts.
- 9 28) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
10 claim 24 in competition assays.
- 11 29) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
12 claim 24 to identify the presence of a T. parva antigen in tissue harvested from an animal.
- 13 30) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
14 claim 24 to quantify the expression of T. parva antigen in cells or biologic fluids.
- 15 31) A diagnostic kit, for the detection of protozoan infection, using the monoclonal antibody or  
16 specific binding parts thereof, that are products of claim 24.
- 17 32) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
18 claim 24 in a Western blot assay.
- 19 33) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
20 claim 24 in an ELISA assay.
- 21 34) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
22 claim 24 in detection of specific molecules using a flow microfluorimeter.
- 23 35) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
24 claim 24 to purify antigen that binds to the antibody or its parts.

- 1 36) The use of the monoclonal antibody or specific binding parts thereof, that are products of  
2 claim 24 in competition assays.
- 3 37) A vector containing an isolated polynucleotide sequence comprising SEQ ID NO: 1, SEQ ID  
4 NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, or SEQ ID NO: 5.
- 5 38) The vector according to claim 39, wherein said vector further comprises a polynucleotide  
6 encoding a polypeptide of interest operably linked to an isolated polynucleotide sequence  
7 comprising SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5,  
8 SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID  
9 NO: 11, SEQ ID NO: 12, or SEQ ID NO: 13.
- 10 39) The vector according to claim 40, wherein said vector further comprises polynucleotide  
11 sequences encoding one or more selectable marker proteins or polypeptides.
- 12 40) The vector of any one of claim 40, 41, 42, or 43, wherein the vector is transfected into a cell  
13 to result in a transformed host cell.
- 14 41) A method for the recombinant production of a polypeptide of interest comprising a vector  
15 comprising an isolated polynucleotide sequence comprising SEQ ID NO: 1, SEQ ID NO: 2,  
16 SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO:  
17 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, or SEQ ID NO: 13,  
18 operably linked to a polynucleotide encoding a polypeptide of interest and expressing said  
19 polypeptide of interest.
- 20 42) The method according to claim 45, wherein said method further comprises the recovery of  
21 said polypeptide of interest.
- 22 47) The method according to claim 45, wherein said vector further comprises polynucleotide  
23 sequences encoding one or more selectable marker proteins or polypeptides.
- 24 48) The method of claim any one of claim 45, 46, or 47, further comprising transfecting the  
25 vector into host cells .

- 1 49) The polynucleotide sequence claimed in any one of claims 10,11, 12, 13 or 14, further  
2 comprising the use of such sequence as a diagnostic tool to detect nucleotide sequence in  
3 animal tissue.
- 4 50) The polynucleotide sequence claimed in any one of claims 10,11, 12, 13 or 14, further  
5 comprising the use of such sequence as a probe of nucleic acids immobilized on a solid  
6 substrate.
- 7 51) The polynucleotide sequence claimed in an one of claims 10,11, 12, 13 or 14, further  
8 comprising the use of such sequence to quantify the expression of mRNAs present in, or  
9 harvested from, animal tissue.
- 10 52) The complimentary strand of the polynucleotide sequence claimed in any one of claims  
11 10,11, 12, 13, or 14.
- 12 53) The use of the complimentary strand of claim 52 as an anti-sense polynucleotide, in either  
13 single stranded or double stranded form
- 14 54) A isolated polypeptide having the following characteristics:  
15 (a) said purified protein or peptide encoded by the expression of cDNA, wherein the purified  
16 protein is recognized by T. parva pathogen-specific cytotoxic lymphocytes.  
17 and  
18 (c) said cDNA comprises;  
19 (i) nucleotide sequence that encodes a surface protein of an intracellular pathogen  
20 (ii) or a nucleotide sequence that encodes a surface protein of an intracellular pathogen that has  
21 been modified or mutated, wherein said encoded protein or peptide stimulates said cytotoxic T  
22 cells.
- 23 52) A isolated polypeptide selected from the group consisting of SEQ ID NO:7, SEQ ID NO:8,  
24 SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12 and SEQ ID NO:13.
- 25 53) The isolated polypeptide according to claim 55, wherein said protein or peptide is detectable  
26 in isolates of protozoa.
- 27 54) The isolated polypeptide according to claim 56, wherein said protein or peptide is detectable  
28 in isolates of Theileria.



- 1 55) The isolated polypeptide according to claim 57, wherein said protein or peptide is detectable  
2 in isolates of *Theileria parva*.
- 3 56) An immunogenic composition comprising the polypeptide according to claim 54 and a  
4 pharmaceutically acceptable carrier.
- 5 57) An immunogenic composition comprising, a cell containing an expressable recombinant  
6 nucleotide molecule, wherein the nucleotide sequence contains one or more of SEQ. ID.  
7 NO:1, SEQ. ID. NO:2, SEQ. ID. NO:3, SEQ. ID. NO:4, and SEQ. ID. NO:5.
- 8 58) An immunogenic composition, as in claim 60, wherein the nucleotide sequence is obtained  
9 from carrying out a PCR amplification of all or a portion of SEQ. ID. NO:1, SEQ. ID. NO:2,  
10 SEQ. ID. NO:3, SEQ. ID. NO:4, and SEQ. ID. NO:5.
- 11 59) The immunogenic composition according to claim 60 or 61, wherein said composition  
12 induces cytotoxic T cells.
- 13 60) A pharmaceutical composition, comprising the polypeptide of any one of claims 55, 56, 57,  
14 or 58 and a pharmaceutically acceptable carrier.
- 15 61) A method for in vitro diagnosis of a protozoan infection which comprises contacting  
16 peripheral blood monocytes taken from an individual with the isolated polypeptide according  
17 to claim 54 or 55 under conditions suitable for binding between said polypeptide and the  
18 cytotoxic T lymphocytes; detecting binding between the purified protein or peptide and the  
19 cytotoxic T lymphocytes; and correlating the binding with the presence of the pathogenic  
20 protozoan.
- 21 62) The method according to claim 64, wherein said method detects the presence of infection by  
22 a plasmodium.
- 23 63) The method according to claim 65, wherein said method detects the presence of infection by  
24 *Theileria*.
- 25 64) The method according to claim 66, wherein said method detects the presence of infection by  
26 *Theileria parva*.
- 27 65) The immunogenic composition of any one of claims 59, 60 or 61, in which the antigen is  
28 administered to an animal to prevent subsequent infection by *Theileria parva*.
- 29 66) The immunogenic composition of claim 22, in which the antigen is administered to an animal  
30 to prevent subsequent infection by *Theileria parva*.

1

2 ABSTRACT:

3

4 This invention relates to DNA, protein, and peptide compositions and methods for  
5 identification of intracellular antigens of pathogens, that stimulate animal cytotoxic  
6 lymphocytes in an antigen specific manner. Such antigens are prime candidates for the  
7 development of vaccines, useful for the prevention of intracellular pathogenic diseases  
8 such as East Coast Fever.

## **ANNEX A**

### **CTL target antigen identification**

#### **1. Materials and methods**

#### **2. Results**

### **1. Materials and Methods**

#### **1.1 Generation of a unidirectional cDNA expression library of *Theileria parva* schizont stage**

##### **1.1.1 Generation of *Theileria parva* Muguga-infected cell lines (TpM)**

Peripheral blood mononuclear cells (PBMC) were isolated from venous blood and infected with *T. parva* (Muguga) sporozoites in accordance with published protocols (Goddeeris & Morrison, 1988). Infected cell lines were maintained in RPMI-1640 supplemented with 10% Fetaclone II (Hyclone, UK; tested for BVDV & *Mycoplasma spp.*), 100IU/ml Penicillin, 100µg/ml Streptomycin, 50µg/ml Gentamycin,  $5 \times 10^{-5}$ M 2-mercaptoethanol and 2mM L-Glutamine and passaged 1/5 three times a week.

##### **1.1.2 Purification of schizonts**

- Harvest TpM in log growth phase, in 50 ml falcon tubes and spin down at 1500 rpm for 10 minutes at room temperature.
- Resuspend the cell pellet in 50 ml of plain culture medium (RPMI 1640 with 25mM HEPES) and count.
- Spin down again to pellet the cells.

- Resuspend the cell pellet in the residual medium.
- Mix cell suspension with diluted antibody (anti-class I; IL-A19; dilution 1:500 or IL-A88, dilution 1:100), diluted in plain culture medium to yield a final concentration of  $1 \times 10^8$  cells per ml.
- Incubate at on ice for 30 minutes.
- Wash of excess antibody by adding plain culture medium to 50 ml and centrifuging at 1000 rpm for 10 minutes at room temperature.
- Resuspend the cell pellet in 2-3 ml of plain culture medium, and add freshly thawed rabbit serum (Complement, previously tested for toxicity) to a final dilution of 1:3 to 1:5 and to a final concentration of  $5 \times 10^7$  or  $1 \times 10^8$  cells per ml.
- Incubate at 37 °C for 30-90 minutes either in a water bath with shaking at 15 minutes interval until lysis is complete (check lysis by trypan blue differential staining using microscope).
- Spin at low speed (800 rpm) for 3 minutes to remove large cell debris.
- Transfer the schizont-rich supernatant into a fresh tube and centrifuge at 2,500 rpm for 10 minutes at 4 °C.
- Discard the supernatant and resuspend the schizont-rich pellet in 6 ml of plain culture medium.
- Separate schizonts from host cell debris and nuclei by Percoll gradient centrifugation as described by Baumgartner et al. (Baumgartner *et al.*, 1999).

### 1.1.3 Isolation of mRNA-poly(A<sup>+</sup>) from schizonts

- Poly(A<sup>+</sup>)RNA was isolated from frozen pellets containing approximately  $3 \times 10^9$  schizonts (from F100 TpM) using FastTrack 2.0 kit (Invitrogen). Frozen schizont pellets were thawed rapidly at room temperature then resuspended and lysed in 15 ml FastTrack 2.0 lysis buffer. The purification of poly(A<sup>+</sup>)RNA was done following the recommendations of the manufacturer (Invitrogen). The total amount of poly(A<sup>+</sup>)RNA isolated and dissolved in 20  $\mu$ l of sterile distilled water, as estimated by ethidium-bromide-stained agarose gel, was 2  $\mu$ g.
- For subsequent analyses, the 20  $\mu$ l poly(A<sup>+</sup>)RNA in a microcentrifuge tube was precipitated on dry-ice for 20 minutes after addition of 2  $\mu$ l of 3 M sodium acetate buffer at pH 5.2, 50  $\mu$ l of absolute ethanol. The frozen solution was thawed and centrifuged at 14,000 rpm for 15 minutes at 4 °C. The ethanol was removed and the RNA precipitate was resuspended in 1 ml of 70% ethanol and re-centrifuged. After removing the ethanol, the RNA was air-dried at room temperature and re-dissolved in 6  $\mu$ l of water. It was then used to construct a unidirectional cDNA library using Invitrogen cDNA synthesis kit (Cat. No. 11917-010).

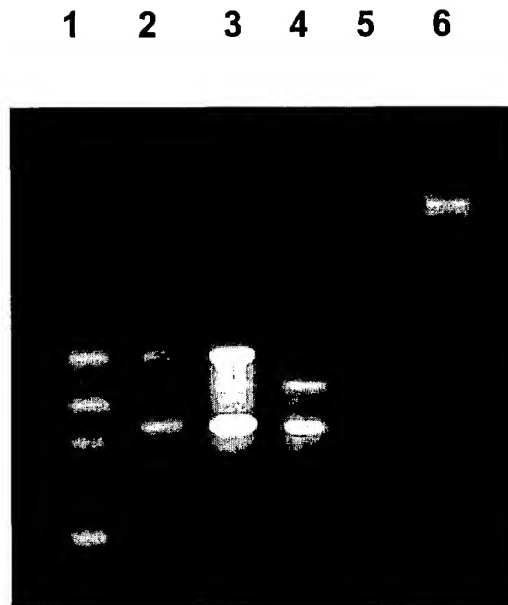


Figure 1: Ethidium bromide-stained agarose gel of RNA. Lane1, GIBCO-BRL RNA size markers; lane2, bovine lymphocyte total RNA; lane3, schizont-infected bovine T-lymphocytes total RNA; lane 4, schizont total RNA; lane 5, schizont poly(A<sup>+</sup>)-RNA; lane 6, schizont poly(A<sup>-</sup>)-RNA.

#### 1.1.4 Synthesis of cDNA

- In a microcentrifuge tube, schizont poly(A<sup>+</sup>)RNA (2 µg in 6 µl water) was mixed with 2 µl (1 µg) of oligo(dT) *Not* I primer-adaptor and incubated at 70 °C for 10 minutes.
- The mixture was spun down briefly then 4 µl 5x first strand buffer, 2 µl 0.1 M DTT, 1 µl 10 mM dNTPs mix and 1 µl [ $\alpha$ -<sup>32</sup>P]dCTP (1 µCi/µl) were added, mixed and incubated at 37 °C for 2 minutes.
- After adding 4µl SuperScript II RT, the reaction was mixed gently and incubated at 37 °C for 1 hour. The reaction was placed on ice.

- One microlitre was removed and used for the calculation of the first strand cDNA yield. The remaining 19  $\mu$ l was used for the second strand cDNA synthesis.
- To the remaining 19  $\mu$ l first strand cDNA reaction placed on ice, 92  $\mu$ l H<sub>2</sub>O, 30  $\mu$ l 5x second strand buffer, 3  $\mu$ l 10 mM dNTP mix, 1  $\mu$ l *E. coli* DNA ligase (10 U/ $\mu$ l), 4  $\mu$ l *E. coli* DNA polymerase I (10 U/ $\mu$ l) and 1  $\mu$ l *E. coli* RNase H (2 U/ $\mu$ l) were added.
- The reaction was mixed gently, spun down briefly and incubated at 16 °C for 2 hours.
- Two  $\mu$ l (10 U) T4 DNA polymerase was added to the reaction, mixed and incubated for additional 5 minutes at 16 °C.
- The reaction was stopped by placing the tube on ice and adding 10  $\mu$ l 0.5 M EDTA.
- One hundred and fifty microlitres of phenol-chloroform solution was added to the cDNA reaction, and then the mixture was vortexed thoroughly for 30 seconds and spun in a microcentrifuge at 14,000 rpm for 10 minutes at 4 °C.
- The top aqueous phase was transferred in a fresh tube and extracted with 150  $\mu$ l chloroform as above.
- The top aqueous phase was transferred in a fresh tube to which 75  $\mu$ l 7.5 M ammonium acetate and 450  $\mu$ l absolute ethanol were added and mixed.
- The tube was immediately centrifuged at 14,000 rpm for 20 minutes at room temperature.

- The supernatant was carefully removed and the cDNA pellet was gently overlaid with 0.5 ml of 70% ethanol. The tube was centrifuged at 14,000 rpm for 2 minutes at room temperature.
- The supernatant was removed and the cDNA pellet was air-dried at room temperature and then redissolved in 18  $\mu$ l H<sub>2</sub>O.

#### 1.1.5 Ligation and fractionation of cDNA

- The ligation was done using Invitrogen kit. The ligation reaction was prepared by adding to the tube containing the 18  $\mu$ l of ds-cDNA the following reagents: 10  $\mu$ l 5x Adapter buffer, 10  $\mu$ l BstX I adapter (at 1  $\mu$ g/ $\mu$ l in H<sub>2</sub>O) 7  $\mu$ l 0.1M DTT and 5  $\mu$ l T4 DNA ligase. The reaction was mix gently and incubated at 16 °C for 20 hours.
- The ligation reaction was stopped by incubating the tube at 70 °C for 10 minutes then chilling on ice.
- To the 50  $\mu$ l of ligation reaction, the following reagents were added and the mixture incubated at 37 °C for 90 minutes: 30  $\mu$ l H<sub>2</sub>O, 10  $\mu$ l 10x buffer 3 (15 U/ $\mu$ l, GIBCO).
- A Sepharose CL-2B gel filtration chromatography was prepared during Not I digestion reaction by washing 4 times with 0.8 ml TEN buffer (TEN buffer = 10 mM Tris-Cl, pH 7.5, 0.1 mM EDTA, 25 mM NaCl). The column is very slow and each wash takes about approx. 20 minutes.
- Fifty microlitre of H<sub>2</sub>O was added to the Not I digestion (final volume: 100  $\mu$ l) and the reaction was loaded onto the column at the center of the resin.



- All the flow through in a single 1.5 ml Eppendorf tube labeled tube 1.
- Then 100  $\mu$ l of TEN was added onto the column and the flow through collected in a single 1.5 ml Eppendorf tube labeled tube 2.
- From this step, TEN was added to the resin in fractions of 100  $\mu$ l and only one drop of the flow through was collected in each tube. Fraction collection was done up to tube 24.
- Each tubes containing cDNA fractions from chromatography was placed in a counting container (with no scintillation fluid) and counted using the tritium channel of a Beckman counter machine.
- Fractions representing the first 300  $\mu$ l and fractions collected after the first 3000  $\mu$ l were discarded.
- The remaining were grouped into 4 fractions represented by tubes 7 to 12 (fraction A), tubes 13 and 14 (fraction B), tubes 15 and 16 (fraction C) and tubes 17 and 18 (fraction D). The amount of cDNAs was as followed: fraction A, 13 ng; fraction B, 46 ng, fraction C, 95 ng; and fraction D, 82 ng.
- Five microlitre of each fraction was run on a 1% agarose gel in 0.5x TAE buffer at 7 V/m. The gel was then incubated in 10% TCA containing 1% NaPPi for 20 minutes.
- The gel was placed on a piece of Whatman 3M paper and overlaid with a pad of absorbent paper and a 500g weight for 10 minutes.
- The pad of absorbent paper and weight were removed and the gel was dried in a gel dryer under vacuum for 20 minutes.
- The dried gel was exposed to autoradiographic film at

–80°C overnight (Figure 2).

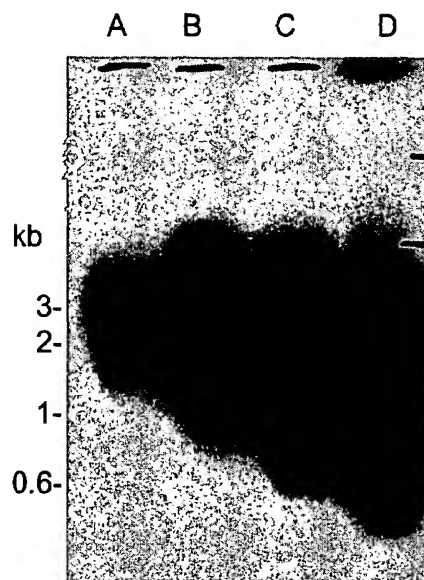


Figure 2: Autoradiograph of electrophoresis gel of the 4 fractions (A, B, C and D) of <sup>32</sup>P-labelled ds-cDNA. The sizes in kilobases are indicated.

#### 1.1.6 Cloning of cDNA into pcDNA3 plasmid vector

- Each of the 4 cDNA fractions was adjusted to 150 µl by adding H<sub>2</sub>O and 1 µl (10 µg yeast t-RNA). Then 75 µl of 7.5 M NH<sub>4</sub>OAc, followed by 450 µl of ice-cold absolute ethanol were added. The mixture was vortexed thoroughly and immediately centrifuged at room temperature for 20 min at 14,000 rpm.
- The supernatant was removed carefully (and disposed properly), and then the pellet was overlaid with 0.5 ml of ice-cold 70% ethanol and centrifuged for 2 min at 14,000 rpm. The supernatant was removed carefully (and disposed properly).

- The cDNA was air-dried at room temperature and re-dissolved in H<sub>2</sub>O as follow: fraction A in 14 µl, fraction B in 28 µl, fraction in 70 µl and fraction D in 56 µl.
- Bacteria *E. coli* DH5α cells containing expression plasmid pcDNA3 were inoculated in 50 ml of 2x YT medium (tryptic soy broth 16 g/l; yeast extract, 10 g/l; NaCl, 5.0 g/l) containing 100 µg/ml ampicillin and grown overnight at 37°C with vigorous shaking.
- Bacteria culture was transferred in a 50 ml falcon centrifuge tube and the cells were pelleted by centrifugation at 3,000 rpm in a minifuge (Heraeus Christ) for 10 minutes at room temperature.
- The supernatant was discarded and plasmid DNA was purified from bacteria pellet using the QIAprep miniprep (Qiagen)) DNA purification system. Plasmid DNA was eluted in H<sub>2</sub>O.
- Twenty microgram of plasmid DNA (8 µl) was digested with 15 µl (10 U/µl) of the restriction enzyme Not I (Promega) in 300 µl reaction mixture containing 1x buffer D (Promega) at 37 °C for 3 hours. Complete digestion and linearization of the plasmid were confirmed by running 2 µg (30 µl) of the plasmid on a 0.7% agarose gel.
- The rest of the Not I reaction mix (270 µl) was mixed with 15 µl of BstX I (10 U/µl) and 15 µl of water and incubated at 55 °C for 3 hours.
- Then 15 µl of the reaction was checked on a 0.7% agarose gel.
- The rest of the Not I/BstX I double digested pcDNA3 vector (285 µl) was dephosphorylated with 3 µl (3 U) alkaline phosphatase at 37 °C for 2 hours.

- Digested and dephosphorylated plasmid vector was purified using CHROMA SPIN-1000 column (Clontech). The concentration of the vector was adjusted at 50 µg/ml in water.
- 1 µl (50 ng) of pcDNA3 was mixed with 1 µl (1 U) T4 DNA ligase (Promega), 1 µl 10x ligation buffer and 7 µl H<sub>2</sub>O and incubated at 16 °C for 16 hours.
- The reaction was stopped by heating at 70 °C for 10 minutes. Then 10 µl of H<sub>2</sub>O is added to give a final concentration of 2.5 ng/µl.
- 40 µl of electro-competent bacteria DH5α previously prepared and tested for their efficiency were electroporated with 2.5 ng (1 µl) of pcDNA3 at 2.49 kV/25 µF using a cuvette of 0.2 cm electrode (BioRad) cold on ice.
- The electroporated bacteria were resuspended in 1 ml LB medium, incubated at 37 °C for 1 hour and then plated on LB agar plates and incubated overnight at 37 °C.
- Plasmid vector and ds-cDNAs were mixed in a ratio of 50 ng of vector for 10-20 ng of cDNAs in a 20-µl ligation reaction containing 4 units of T4 DNA ligase (Promega). Thus the final volume of ligation of different was as follow: 20 µl for the 13 ng of fraction A, 40 µl for the 28 ng of fraction B, 100 µl for the 95 ng of fraction C, and 80 µl for the 82 ng of fraction D. Fraction E was a control ligation reaction of 40 µl containing 40 ng pcDNA3 and no ds-cDNAs.
- The ligation reactions were incubated at 4 °C for 16 hours.
- Water was added to each ligation mix to a final volume of 150 µl. Followed by 1 µl of 10 µg/µl yeast t-RNA.

- Each ligation was extracted with 150  $\mu$ l phenol-chloroform by vortexing for 30 seconds and spinning at 14,000 rpm for 10 minutes at 4 °C.
- The top aqueous phase was re-extracted with 150  $\mu$ l chloroform.
- The top aqueous phase was transferred in a fresh tube and the DNA was precipitated by adding 0.1 volume of 3 M NaAc, pH 5.2, and 2.5 volumes (375  $\mu$ l) of ice-cold absolute ethanol, followed by mixing briefly and incubating for 15 minutes on dry-ice.
- The tubes were thaw at room temperature and centrifuged at 14,000 rpm for 10 minutes at 4 °C to pellet the DNA.
- The DNA pellet was washed with 1 ml of 70% ethanol, followed by centrifugation at 14,000 rpm for 5 minutes at 4 °C.
- The supernatant was removed and the DNA pellet was air-dried at Room temperature and re-dissolved in H<sub>2</sub>O as follow: fraction A, 5  $\mu$ l; fraction B, 10  $\mu$ l; fraction C, 25  $\mu$ l; fraction D, 20  $\mu$ l; and fraction E, 10  $\mu$ l.
- Frozen electrocompetent DH5 $\alpha$  *E. coli* were thawed on ice and maintained on ice with the cuvettes. In ice-cooled Eppendorf tubes, 2.5  $\mu$ l of DNA was mixed with 40  $\mu$ l of DH5a and transferred in 0.2 cm electrode cuvette for electroporation at 2.49 kV/25  $\mu$ F. Each electroporation was recovered with 1 ml LB and transferred in 12 ml Sterilin tubes, followed by incubation at 37 °C for 1 hour on a shaker at 220 rpm.
- For fraction A, 2 electroporations were done; 4 for fraction B, 10 for fraction C, 8 for fraction D, and only 1 for the control fraction E. Electroporations from the same fraction were pooled together in one tube, mixed and plated.

- For plating, 20  $\mu$ l of each pool was transferred in a fresh tube containing 100  $\mu$ l of LB medium: dilution 1. 10  $\mu$ l of dilution 1 was transferred in another fresh tube containing 100  $\mu$ l of LB: dilution 2; 10  $\mu$ l of dilution 2 was transferred in another fresh tube containing 100  $\mu$ l of LB: dilution 3. Dilutions 1, 2 and 3 were plated onto LB agar plate containing ampicillin at 50  $\mu$ g/ml for the titration of the library. The rest of each electroporation was plated at 2 ml per LB agar plate of 24x24 cm containing ampicillin at 50  $\mu$ g/ml. Plates were incubated overnight at 37°C. Colonies from cultures plated out at low dilutions were counted and the total number of colonies in each fraction calculated (Table 1).

<b>Fractions</b>	<b>Number of independent colonies</b>
A	200
B	12,800
C	355,000
D	298,000
<b>Total library</b>	<b>666,000</b>

Table 1: Titration of cDNA library. The number of independent colonies in each fraction.

- The content of a small tube (50 µl of bacteria) was resuspended in 950 µl of LB medium, mixed, diluted 10 fold several times and plated on agar plates containing ampicillin.
- After incubation overnight at 37 °C, colonies formed on agar plates were counted and the total number of bacteria in each stabilates was determined.
- The quality of the library was assessed as described in Table 2.

	Fraction A	Fraction B	Fraction C	Fraction D
% of colonies containing inserts	100	100	95	85
Size of inserts (kb)	1 - 3	0.4 - 3	0.6 - 2	0.2 – 1.5
% of inserts of <i>T. parva</i> origin*	100	100	100	100

Table 2. Analysis of randomly selected cDNA from each fraction. Plasmid DNA was isolated from 20 cDNA clones from each of the 4 fractions, cut with restriction enzymes Hind III and Xba I and analysed for presence and size of inserts, whose sequences were blasted against non-redundant GenBank and TIGR *T. parva* genome databases (<http://www.tigr.org/tdb/e2k1/tpa1/>). \*An additional 120 cDNA were sequenced and found to be of *T. parva* origin.

#### 1.1.7 Preparation of cDNA pools for immunoscreening with CTL

- Pools of 50 as well as pools of 10 bacteria was prepared from fractions B and C followed by plasmid minipreps purification.

- Aliquots of frozen bacteria in small tubes containing 50  $\mu$ l of bacteria were diluted and plated on agar plates to a density of 40 - 60 colonies per 4 - 6  $\text{cm}^2$  and grown overnight at 37°C.
- Colonies were scraped from the agar plate in pools of about 50 or 10 colonies and resuspended in 3 – 4 ml LB medium with Ampicillin. Bacteria suspensions were cultivated at 37°C for 3 - 4 hrs with shaking.
- Then 0.8 ml of each culture was mixed with 0.8 ml LB medium containing 40% glycerol and frozen as a glycerol at –80°C while the remaining was used for plasmid purification. Plasmid DNA was eluted in 200  $\mu$ l of water then transferred in a well of a 96 wells microwells plate and stored at –20°C until use
- For the transfection each pool of cDNAs was adjusted at 100 ng/ $\mu$ l.

#### 1.1.8 Strategy for resolving positive pools to identify cDNA coding for CTL target antigen

- Competent *E. coli* DH5 $\alpha$  (200  $\mu$ l) with 1  $\mu$ l (100 ng) of DNA from the positive pool
- The transformation is diluted and plated on agar plates to a density of 2000 colonies per 150  $\text{cm}^2$  and grown overnight at 37°C.
- For a positive pool of 10, only 48 individual colonies are picked and grown overnight in 2 ml of 2x YT medium containing ampicillin. Plasmid DNA is then isolated from each bacteria culture by minipreps. Plasmid DNA are eluted in 200  $\mu$ l of water and then transferred in a well of a 96 wells microwells plate



and stored at  $-20^{\circ}\text{C}$ . The working concentration is adjusted at  $100\text{ ng}/\mu\text{l}$  and stored at  $-20^{\circ}\text{C}$  until use.

- For a positive pool of 50, a three-way matrix was used as follow: 256 individual colonies are picked and seeded in  $200\text{ }\mu\text{l}$  of 2x YT medium containing ampicillin in 4 flat- bottom microwells plates. This gives 64 colonies per plate.
- Bacteria were grown at  $37^{\circ}\text{C}$  overnight.
- From each plate containing 64 cultures, 12 pools of bacteria, each containing 16 individual bacteria cultures, are prepared by pooling  $50\text{ }\mu\text{l}$  of each individual bacteria culture as shown in the table 7 below. In total, 48 pools of 16 individual bacteria are generated. Add glycerol to 20% in the each of the remaining cultures in Table 3 and store at  $-70^{\circ}\text{C}$  until the result of the screening to pick the positive colonies.

### **Use of a 3-way matrix for rapid recovery of a positive cDNA clone**

Competent *E. coli* cells ( $200\text{ }\mu\text{l}$ ) were transformed with  $1\text{ }\mu\text{l}$  ( $100\text{ ng}$ ) of DNA from the positive pool of 50 cDNAs. Transformed *E. coli* were diluted and plated on agar plates to a density of 2000 colonies per  $150\text{ cm}^2$  and grown overnight at  $37^{\circ}\text{C}$ . Two hundred fifty six (approximately 5x coverage of 50 cDNA per positive pool) individual colonies were picked and seeded in  $200\text{ }\mu\text{l}$  of 2x YT medium containing ampicillin in 4 flat-bottom microwells plates to have 64 single bacteria colony cultures per plate arrayed in a  $8 \times 8$  format and numbered from 1 to 64 as tabulated below (this is one of 4 plates containing 64 *E. coli* cultures).

		1	9	17	25	33	41	49	57		
		2	10	18	26	34	42	50	58		
		3	11	19	27	35	43	51	59		
		4	12	20	28	36	44	52	60		
		5	13	21	29	37	45	53	61		
		6	14	22	30	38	46	54	62		
		7	15	23	31	38	47	55	63		
		8	16	24	32	40	48	56	64		

Twelve pools of 16 individual cultures were generated from the 4 plates as shown in the table below. This format is high throughput in that multi-channel pipetting accelerates the process. The 3-way matrix is such that each individual bacteria culture is present in 3 pools. Pools of bacteria cultures were grown at 37°C for 3 hours with shaking before plasmid DNA was extracted from pooled 16 aliquots of each pool. Glycerol was added to the remaining cultures in microplates and stored at –70°C. Plasmid DNA was adjusted at 100 ng/μl and used for transfection in screening assays. A positive cDNA clone is detected in 3 pools. For example, if colony #25 contained the target cDNA clone, pools #2, 5 and 11 would be detected positive. Plasmid DNA is then prepared from the glycerol stock of bacteria colony #25 from the 3 pools individually and re-tested and sequenced.

	P I	P ol	P ol	Pool	Po I	Po I	Po I	Pool	P ol	P ol	P ol	Pool
	1	2	3	4	5	6	7	8	9	10	11	12
Colony #	1	17	33	49	1	2	3	4	1	5	9	13
	2	18	34	50	5	6	7	8	2	6	10	14
	3	19	35	51	9	10	11	12	3	7	11	15
	4	20	36	52	13	14	15	16	4	8	12	16
	5	21	37	53	17	18	19	20	17	21	25	29
	6	22	38	54	21	22	23	24	18	22	26	30
	7	23	39	55	25	26	27	28	19	23	27	31
	8	24	40	56	29	30	31	32	20	24	28	32
	9	25	41	57	33	34	35	36	33	37	41	45
	10	26	42	58	37	38	39	40	34	38	42	46
	11	27	43	59	41	42	43	44	35	39	43	47
	12	28	44	60	45	46	47	48	36	40	44	48
	13	29	45	61	49	50	51	52	49	53	57	61
	14	30	46	62	53	54	55	56	50	54	58	62
	15	31	47	63	57	58	59	60	51	55	59	63
	16	32	48	64	61	62	63	64	52	56	60	64

Table 3: A 3-way matrix for rapid resolution of positive pools. In this format each colony is represented three times.

- Pools of bacteria are grown at 37 °C for 3 hours with shaking before they are used for plasmid purification by minipreps. Plasmid DNA are eluted in 200 µl of water and then transferred in a well of a 96 wells microwells plate and stored at –20 °C.

- The working concentration is adjusted at 100 ng/μl and stored at –20 °C until use.

## 1.2 Immunisation & details of cattle

The 7 cattle used in this invention were pure bred *Bos indicus* (Boran) pure bred *Bos taurus* (Friesian) or crossbreeds. Cattle were immunised by ‘infection and treatment’ against the Muguga stock of *T. parva* by simultaneous inoculation of sporozoites and long-acting oxytetracycline at 20mg/kg BW (Radley et al 1975). Cryopreserved sporozoites (Stabilate # 4133) were thawed and diluted 1/20 as previously described. Animals were given a subcutaneous injection of 1ml of diluted sporozoites 2cm above the right parotid lymph node. Animals were monitored daily for changes in rectal temperature and from day 5 post challenge lymph node biopsies were taken using a 21G needle. Giemsa stained biopsy smears were examined for the presence of schizont infected cells and scored on a scale of 1-3. Animals suffering from moderate reactions were treated with Buparvaquone (Butalex, Mallinckrodt Veterinary Ltd, UK).

## 1.3 Generation, characterisation and maintenance of CTL

### 1.3.1 Establishment of TpM

Prior to immunisation, venous blood was collected from the eleven animals, PBMC were purified and infected *in vitro* with *T. parva* (Muguga) sporozoites (Goddeeris & Morrison, 1988).

### 1.3.2 Generation of *T. parva* specific CD8<sup>+</sup> CTL

Venous blood from immunised animals was collected from 4 weeks post-immunisation. PBMC were prepared as described previously (Goddeeris & Morrison, 1988). PBMC were adjusted to  $4 \times 10^6$ /ml in CTL medium (RPMI-1640 without HEPES supplemented with 10% FBS (HyClone; tested for BVDV & mycoplasma spp.), L-glutamine, 2-mercaptoethanol and antibiotics as described above) and 1ml/well added to 24 well plates (Costar, Corning, NY, USA). PBMC were co-cultured with irradiated (50Gy) autologous *T. parva* infected cells (TpM) at  $2 \times 10^5$ /well and incubated for 7 days at 37°C in a 5% CO<sub>2</sub> humidified atmosphere. Cells were harvested by aspiration and dead cells removed by centrifugation over Ficoll-Paque Plus (Amersham Pharmacia Biotech, Uppsala, Sweden). After washing in CTL medium, cells were added to 24 well plates ( $3 \times 10^6$ /well) and co-cultured with irradiated PBMC (filler cells) at  $1 \times 10^6$ /well and TpM at  $2 \times 10^5$ /well for 7 days as before. Viable cells were harvested as described above, adjusted to  $2 \times 10^6$ /well and stimulated with  $4 \times 10^5$ /well irradiated TpM and  $2 \times 10^6$ /well irradiated autologous PBMC as filler cells.

<sup>51</sup>Chromium Release Assay: Autologous and allogeneic TpM in log phase of growth were resuspended at  $2 \times 10^7$ /ml cytotoxicity medium (RPMI-1640 medium with 5% Fetaclone II). 100µl of the target cells were mixed with 100µl (100µCi) of <sup>51</sup>Cr-sodium chromate and incubated for 1 hour at 37°C. Cells were washed 3 times in 7ml of cytotoxicity medium by centrifugation at 1500 rpm for 7 min at RT and resuspended at  $1 \times 10^6$ /ml. Viable cells were harvested from TpM stimulated

lines 7 days post-stimulation (effector cells) and resuspended in cytotoxicity medium at  $2 \times 10^7$ /ml. Two-fold doubling dilutions of effector cells were distributed in duplicate ( $100 \mu\text{l}$ /well) to 96-well half area (A/2) flat-bottom culture plates (Costar, Corning, NY, USA) resulting in a range of effector cell concentrations of  $4 \times 10^6$  to  $2.5 \times 10^5$ /well. Target cells were added to each well containing effector cells ( $50 \mu\text{l}$ /well) resulting in target cell ratios ranging from 80:1 to 5:1. In separate triplicate wells target cells were added to  $100 \mu\text{l}$  cytotoxicity medium or 1% Tween20 to measure spontaneous and maximal release of the label respectively. Plates were incubated for 4 hours at  $37^\circ\text{C}$  in 5%  $\text{CO}_2$  humidified atmosphere. Cells were resuspended in wells by repeated pipetting and pelleted by centrifugation at  $180 \times g$  at room temperature.  $75 \mu\text{l}$  of supernatant was transferred from each well into sample vials (Milian, Geneva, Switzerland) and gamma emissions counted in a gamma counter (Micromedic MEplus, TiterTek, Huntsville, AL, USA). Results were calculated and expressed as percent cytotoxicity ( $= 100 \times (\text{test release} - \text{spontaneous release}) / (\text{maximum release} - \text{spontaneous release})$ ).

Evidence for MHC class I restricted lysis was assessed by the capacity of monoclonal antibodies recognising bovine MHC class I to inhibit lysis. Cells were prepared for the cytotoxicity assay as described above except that target cells were resuspended at double the density ( $2 \times 10^6$ /ml) and  $25 \mu\text{l}$  added first to the plate.  $25 \mu\text{l}$  of either cytotoxicity medium or monoclonal antibody (mAb; IL-A88 diluted 1/15 in cytotoxicity medium was added to target cells and incubated for

30min at room temperature. Serial dilutions of effector cells were then added as described above.

Viable cells from TpM stimulated cultures were harvested day 7 post-infection and CD8<sup>+</sup> T cells were isolated either by positive selection using flow cytometry (FACStar Plus, BD Biosciences, San Jose, CA, USA) or negative selection using Dynabeads (DynaL Biotech, Bromborough, UK).

Positive selection: Cells were adjusted to  $2 \times 10^7$ /ml in sterile monoclonal antibody IL-A105 (specific for bovine CD8) diluted 1/100 and incubated for 30 minutes at 4°C. Cells were washed twice in cold culture medium and resuspended at  $2 \times 10^7$ /ml in sterile goat anti-mouse polyvalent immunoglobulins conjugated to FITC (Sigma). Cells were washed twice in cold medium before being run through a FACStar plus cell-sorter and CD8<sup>+</sup>FITC<sup>+</sup> cells collected.

Negative selection: Cells were adjusted to  $2 \times 10^7$ /ml in sterile mAbs IL-A12 (specific for bovine CD4) and GB21A (specific for bovine  $\gamma\delta$  TCR) diluted 1/100 and incubated for 30 minutes at 4°C. Cells were washed twice in cold PBS and resuspended  $1.4 \times 10^7$ /ml in sterile PBS containing washed sheep anti-mouse IgG Dynabeads according to the manufacturers instructions Cells and beads were rotated for 30 min at 4°C and Dynabead rosetted cells were collected by placing the sample tube in a Dynal Magnetic Particle Concentrator (MPC) for 5 min, the supernatant was removed, transferred to another sample tube and residual Dynabead rosetted cells removed by another incubation in a Dynal MPC. The supernatant was removed and cells washed twice in complete medium.

Cloning: CD8<sup>+</sup> T cells enriched by either of the methods described above were adjusted to 30, 10, and 3 cell/ml and distributed into 96-well, round bottom culture plates containing  $2 \times 10^4$  irradiated autologous TpM,  $5 \times 10^3$  irradiated autologous PBMC and 5U/ml recombinant human IL-2 (HuIL-2; Sigma, Poole, UK) in a final volume of 200 $\mu$ l/well. Plates were incubated at 37°C in humidified incubator containing 5% CO<sub>2</sub> in air. Wells showing significant cell growth were selected for analysis of lysis of autologous TpM in an (<sup>111</sup>In) Indium oxine release assay. The <sup>111</sup>In release assay was performed as described above for <sup>51</sup>Cr release assay except that targets cells were labelled by addition of 5 $\mu$ Ci <sup>111</sup>In/1 x 10<sup>6</sup> cells and incubation for 15 minutes at 37°C. Cells were washed five times with cytotoxicity medium and resuspended at 1x10<sup>5</sup>/ml and added 50 $\mu$ l/well to 96-well V-bottom 96 well plates. 100 $\mu$ l cells from wells showing significant growth were transferred to 96-well V-bottom culture plates (Greiner) and centrifuged (180xg) for 5 min. Cells were resuspended 100 $\mu$ l/well in cytotoxicity medium and transferred to wells containing labelled target cells. Plates were centrifuged as described above to pellet cells before being incubated for 4 hours at 37°C in 5% CO<sub>2</sub> humidified atmosphere.

CTL populations that exhibited lytic activity on autologous infected cells and originated from cell dilutions that gave rise to cell growth in less than 30% of the wells, as they have high probability of being clones, were selected for expansion. The remaining cells were harvested from each well and resuspend them in culture medium (without HEPES) at a concentration estimated to be between 500



and 5000 cells/ml. 100 $\mu$ l of cell suspension was distributed into 96 well, round bottom culture plates. 100 $\mu$ l of autologous irradiated TpM at  $5 \times 10^4$  autologous irradiated TpM in medium containing 5U/ml Hull-2. Between day 14 and 21 post-stimulation clones and polyclonal CTL lines were subcultured in 96 well, round bottom culture plates by co-culturing 5000 CTL/well with 25,000 autologous irradiated TpM and 5U/ml Hull-2 in a final volume of 200 $\mu$ l/well.

#### 1.4 Establishment and maintenance of bovine skin fibroblasts

A skin biopsy was taken aseptically from the ears of cattle and placed in a 50ml falcon tube containing Alsever's solution. In the laboratory laminar flow hood, the biopsy was placed into sterile petri dishes (Sterilin) containing 1ml of 0.25% Trypsin-EDTA. Using a sterile scalpel blade, the sample was cut into small pieces and placed into a 50 ml falcon tube containing Trypsin-EDTA. The preparation was placed in a shaking incubator for 1-1.5 hr at 37°C with gentle continuous shaking. This facilitates detachment of cells. This digestion was stopped by addition of 1ml heat-inactivated FBS. The cell suspension was centrifuged for 10 min at 200 $\times$ g and the cell pellet re-suspended in 6 ml of Dulbecco's minimum essential medium, DMEM (Gibco-BRL, Paisley, UK) supplemented with 10% heat-inactivated fetal bovine serum (FBS) (Hyclone, Logan, UT), 400 IU/ml penicillin, 300  $\mu$ g/ml streptomycin and 2 mM L-glutamine. This cell suspension was seeded in 5ml amounts into 25-cm<sup>2</sup> flasks and incubated at 37°C in a 5% CO<sub>2</sub> humidified atmosphere. Cultures were examined

microscopically after every 4 days for growth and half the medium was replaced with fresh one until growing colonies were evident. Once positive colonies were identified they were rinsed with  $\text{Ca}^{2+}$ - and  $\text{Mg}^{2+}$ -free PBS/EDTA (0.02% EDTA) and detached by 2 min incubation at 37°C with Trypsin-EDTA solution containing 2.5 mg/ml Trypsin and 0.2 mg/ml EDTA in HBSS (Sigma). Following a wash in complete DMEM containing 10% FBS the skin fibroblasts (SF) were passaged into 25-cm<sup>2</sup> tissue culture flasks (Costar). Cells were maintained in complete DMEM containing 10% FBS, 200 IU/ml penicillin, 100 µg/ml streptomycin and 2 mM L-glutamine and passaged every 3 days at a ratio of 1:3. Cells were expanded in 75-cm<sup>2</sup> tissue culture flasks (Costar) until confluent yielding  $3\text{--}4 \times 10^6$  cells/flask. For consistency a liquid nitrogen cell bank was prepared with  $2 \times 10^6$  cells/freezing vial in 10% DMSO in FBS.

### **1.5 Establishment and maintenance of bovine testicular endothelial cells**

Bovine testicular vein or pulmonary artery EC lines were established as described by Byrom and Yunker (1990) with the modifications of Mwangi et al (1998). Briefly, the vein was placed in wash buffer consisting of  $\text{Ca}^{2+}$ - and  $\text{Mg}^{2+}$ -free PBS supplemented with 400 IU/ml penicillin, 400 µg/ml streptomycin and 5 µg/ml fungizone. The vessel was then slit longitudinally, washed twice before being cut into 1-cm<sup>2</sup> pieces and then placed lumen side down on a drop of collagenase (1 mg/ml) and incubated for 1 h at 37°C. The cell suspension was centrifuged for 5 min at 200×g and re-suspended in 24 ml 2× complete DMEM (Gibco-BRL, Paisley, UK) supplemented with 20% heat-inactivated fetal bovine

serum (FBS) (Hyclone, Logan, UT), 400 IU/ml penicillin, 300 µg/ml streptomycin, 5 mg/ml fungizone, 300 µg/ml endothelial growth supplement (Sigma, St Louis, MO) and 2 mM L-glutamine. One ml of the cell suspension was seeded in each well of a 24-well tissue culture plate (Costar, Cambridge, MA, USA). Once confluent, monolayers in each well were rinsed with  $\text{Ca}^{2+}$ - and  $\text{Mg}^{2+}$ -free PBS/EDTA (0.02% EDTA) and detached by 2 min incubation at 37°C with 0.25% trypsin/EDTA solution containing 2.5mg/ml trypsin and 0.2 mg/ml EDTA in HBSS (Sigma). Following a wash in DMEM containing 10% FBS the cells were passaged into 25-cm<sup>2</sup> tissue culture flasks (Costar). Cells were maintained in complete DMEM containing 10% FBS, 200 IU/ml penicillin, 100 µg/ml streptomycin, 2.5 mg/ml fungizone, 2 mM L-glutamine and passaged every 3 days at a ratio of 1:3. Cells were expanded in 75-cm<sup>2</sup> tissue culture flasks (Costar) until confluent yielding between  $3 \times 10^6$  and  $4 \times 10^6$  cells per flask. For consistency a liquid nitrogen cell bank was prepared with  $2 \times 10^6$  cells per freezing vial. Cells were raised from nitrogen into 75-cm<sup>2</sup> flasks and used between the fourth and tenth passages.

### **1.6 Immortalization of bovine skin fibroblasts and endothelial cells**

Cells were immortalized with the SV40 early region gene by transfection of an expression plasmid psvNeo (ATCC code 37150). The product is supplied as a freeze dried to which 300ul of 2XYT with ampicillin was added and fully resuspended. The 300ul suspension was sub-cultured into 6mls and divided into 3mls each. The cultures were incubated overnight at 37°C and then subcultured

again into 50mls of 2XYT with ampicillin overnight at 37°C. Maxiprep of psvNeo plasmid were made for use in immortalizations. Plasmid DNA for transfection was standardized to 2mg/ml.

The DNA mix was prepared by mixing 5µl DNA with 495µl of DMEM with antibiotics but no serum. A working dilution of Fugene 6 transfection reagent (Roche) was prepared by mixing 15µl of Fugene with 485µl of DMEM. Mixing was done using the 2ml sterile non-pyrogenic, DNase and RNase free cryopreservation tubes. The diluted Fugene was added to the diluted DNA drop wise with constant tapping at the end of the tube. The Fugene-DNA complex was allowed to form at room temperature for 30 min.

Cells were cultured in 6-well plates (Costar) at a density of  $2 \times 10^5$  cells per well and grown to confluence overnight at 37°C, 5% CO<sub>2</sub> in a humidified incubator. Culture medium was removed completely from the monolayers to be transfected. The Fugene-DNA complex was then added gently onto the monolayer. Plates were incubated at 37°C, 5% CO<sub>2</sub> in a humidified atmosphere for 3-4 hours. The transfection complex was removed; fresh DMEM medium containing 10% fetal calf serum was added and then cultured for a further 72h. Cells from each well were rinsed with PBS/EDTA, detached with Trypsin/EDTA, washed and re-suspended in complete DMEM and sub-cultured into one T-25 flask (Costar). After incubation for 2-3h, normal DMEM medium was removed and replaced with a selection DMEM medium containing 10% FCS and 0.5µg/ml of G418 (2.5µg/ml G418 for endothelial cells) and incubated further. Upon observation of high death rate of cells, half the medium was replaced with fresh selection medium until

growing colonies were evident. This process took 3-4 weeks depending on the cell lines. Positive colonies were sub-cultured into 24 well plates and then to T-25 flasks. Immortalization was confirmed by checking for expression of large T-antigen using an anti-SV40 antibody conjugated to HRP. Further expansion and maintenance of the cells was carried out using complete DMEM.

### **1.7 Transfection of COS-7 cells and immortalised skin fibroblasts with schizont cDNA library**

COS-7 cells and iSF were maintained in T75 and T150 TC flasks with DMEM supplemented with 10% FCS, 2mM L-glutamine and antibiotics (TC medium) as described above. Cells are split 1:4 every 3 days. The day prior to transfection, cells were harvested by the removal of medium, washing in PBS and incubation in 0.25% Trypsin-EDTA for 5 min at 37°C. Once cells had detached TC medium was added and cells removed. Cells were washed by centrifugation at 1200 rpm for 10min and resuspended in TC medium. A viable cell count taken, density adjusted to  $2.0 \times 10^5$ /ml, cells dispensed, 100 $\mu$ l/well, into 96 well flat-bottom TC plates and incubated for overnight at 37°C in a CO<sub>2</sub> (5%) humidified incubator.

DNA was prepared for either single transfections of SF or double transfections (co-transfection of schizont cDNA and BoLA class I cDNA) of COS cells. 6 $\mu$ l of schizont cDNA and 6 $\mu$ l of BoLA class I cDNA (for co-transfection) at 50ng/ $\mu$ l in dH<sub>2</sub>O were added to 150 $\mu$ l unsupplemented DMEM in wells of a 96-well round-bottom plate. FuGENE 6 transfection reagent was pre-warmed to 37°C. 0.9 $\mu$ l or 0.45 $\mu$ l FuGENE 6 was added to each well for double and single transfections

respectively. The well contents mixed by shaking on a Dynatech Varishaker for 1 min and incubated at RT for 20 min. The medium from the 96 well plates containing adherent COS cells and SF was removed and each transfection complex added to triplicate wells (50 $\mu$ l/well). The cells were then incubated for 4 hours at 37°C in a CO<sub>2</sub> (5%) humidified incubator. The transfection complex was removed and replaced with 200 $\mu$ l/well DMEM supplemented as described above and incubate for 24 or 48 hours at 37°C in a CO<sub>2</sub> (5%) humidified incubator.

### **1.8 Detection of CTL recognition of transfected SF and COS-7 cells by IFN- $\gamma$ ELISpot**

24 hours post-transfection, medium was removed from wells containing transfected cells, cells were washed cells with PBS (200 $\mu$ l/well) and detached by the addition of 100 $\mu$ l/well Trypsin-EDTA as described above. Once the cells had detached the contents of each well were transferred to a 96 well round-bottom plates containing 100 $\mu$ l/well cold RPMI with no HEPES supplemented with 10% FCS (CTL medium). The cells were centrifuged at 1200 rpm for 3 min; supernatant removed and resuspended 50 $\mu$ l/well in CTL medium

Schizont specific CTL, generated and maintained as described above, were harvested 7-14 days post-stimulation, transferred to polycarbonate tubes, pelleted at 1200 rpm for 10min and resuspended at 2x10<sup>5</sup>/ml in CTL medium supplemented with 5U/ml HuL-2 (Sigma).

ELISpot plates (Millipore corporation, Bedford, MA, USA) were coated 50 $\mu$ l/well with 2 $\mu$ g/ml of murine anti-bovine IFN- $\gamma$  mAb (CC302; Serotec, UK) and incubated overnight at 4°C. Wells were washed twice with unsupplemented RPMI-1640 and blocked 200 $\mu$ l/well with RPMI-1640 supplemented with 10% FBS by incubating at 37°C for 2 hours. The blocking medium was removed and replaced with 50 $\mu$ l/well CTL and 100 $\mu$ l/well transfected cells. As a positive control, irradiated TpM are serially diluted in COS cells or SF with each at a density of 4x10<sup>5</sup>/ml. and populations containing 32, 16, 8, 4, 2, 1% TpM are added 50 $\mu$ l/well to wells containing CTL. Plates were incubated in a humidified incubator at 37°C for 20 hours. After incubation, the contents of wells were removed and wells washed four times with sterile distilled water supplemented with 0.05% Tween 20 per well and the plate shaken on a shaker for 30 seconds between washes. The process was repeated an additional four times, using PBS supplemented with Tween 20 (PBS-T). Wells were then incubated with 100 $\mu$ l/well rabbit anti-bovine IFN- $\gamma$  antisera diluted 1/1500 in PBS-T supplemented with 0.2% BSA (PBS/BSA) for 1 hour at room temperature. Wells were washed 4 times with PBS-T before being incubated for 1 hour at room temperature with 100 $\mu$ l/well murine monoclonal anti-rabbit IgG conjugated to alkaline phosphatase (Sigma) diluted 1/2000 in PBS-T/BSA. Sigma Fast BCIP/NBT buffered substrate (Sigma) was by dissolving 1 tablet/10ml dH<sub>2</sub>O and passing it through a 0.2 $\mu$ m filter. Plates were washed six times as described above with PBS-T, 100 $\mu$ l/well BCIP/NBT substrate added and plates incubated for 10minutes at room temperature in the dark. The substrate was then removed, wells washed with

copious amounts of H<sub>2</sub>O and plates air-dried at room temperature in the dark. Plates were finally read on an automated ELISpot reader (AID Diagnostica, Strasberg, Germany).

### **1.9 Bioassay for CTL activation based on rapid induction of class II MHC expression by constitutively negative bovine endothelial cells**

Endothelial cells were detached from confluent 75-cm<sup>2</sup> tissue culture flasks by treatment with trypsin/EDTA as described.  $2.5 \times 10^4$  cells were seeded into each well of a 48-well plate in 1 ml of complete DMEM and incubated overnight. Cell-free test supernatants derived from either co-culture of *T. parva*-specific CD8<sup>+</sup> T cell lines and SF transfected with test genes or recombinant bovine IFN- $\gamma$  (rBoIFN- $\gamma$ , the kind gift of Ciba-Geigy, Basel, Switzerland) were dispensed in duplicate wells in a final volume of 160  $\mu$ l per well.

Following 48 h incubation at 37°C, culture supernatants were discarded and the monolayers washed three times in FACS medium (RPMI 1640 supplemented with 2%  $\gamma$ -globulin-free horse serum (Gibco-BRL) and 0.1% sodium azide). Endothelial cells were labeled on ice for 30 min with 100  $\mu$ l of an antibody cocktail comprising equal volumes of bovine class II MHC-specific monoclonal antibodies (mAbs) J11, R1 and IL-A21 (each at a 1/500 dilution of ascitic fluid). After three washes in FACS medium, 100  $\mu$ l of FITC-conjugated goat anti-mouse Ig (Sigma) diluted 1/200 in FACS medium were added. After incubation at 4°C for 30 min and a further three washes in FACS medium, cells were detached as



described previously and cell surface class II MHC expression determined by flow cytometry, using a FACScan (Becton-Dickinson, Sunnyvale, CA, USA). The percentage of class II MHC-expressing cells was determined by comparison with unstimulated EC labeled in an identical manner.

*T. parva*-specific CD8<sup>+</sup> CTL lines were generated and maintained using methods initially described. Test supernatants were collected from 96-well flat-bottomed microtitre plates (Costar) 48 hours after restimulation of resting T cell lines ( $2 \times 10^4$  per well) with  $4 \times 10^5$  COS-7 cells co-transfected with the KN104 gene and test gene(s) or confluent autologous SF transfected with the test gene(s) in a final volume of 200  $\mu$ l for 24 h. These tests were set-up at least in duplicates. Where indicated, class I MHC was blocked using a specific antibody (IL-A88) to check for MHC class I restriction of the CTL lines. Additional negative control supernatants were derived from co-culture of CTL with untransfected immortalized SF or COS-7 cells. Positive control supernatants were obtained from co-culture of CTL with varying proportions of irradiated autologous TpM.

#### **1.10 Detection of CTL lysis of transfected iSF and COS-7 cells**

Autologous iSF or COS-7 cells were seeded in 6-well plates (Costar) at a density of  $2.5 \times 10^5$  /well and incubated for 2 hours at 37°C to allow cells to adhere. For single transfections of iSF, 2 $\mu$ g of test or control cDNA was added to 1ml unsupplemented DMEM containing 3 $\mu$ l FuGENE 6 transfection reagent and incubated for 40min. For COS-7 cells, 2 $\mu$ g of test or control cDNA was added with 2 $\mu$ g of BoLA class I cDNA to 1ml DMEM containing 6 $\mu$ l FuGENE. 1ml of

DNA/Fugene complex was added per well of adherent COS-7 or iSF. Plates were incubated for 4 hours at 37°C, the transfection complex was removed and replaced with 2ml/well of complete DMEM and the plates incubated for a further 20 hours at 37°C. Transfected cells were harvested by removal of medium, washing in PBS, detachment by Trypsin-EDTA and washing in complete DMEM. Transfected cells and TpM were labelled with <sup>51</sup>Chromium and the ability of schizont-specific CTL lines (day 6-8 post-stimulation) to lyse these targets was assessed as described above.

#### **1.11 Use of Exonuclease III digestion to map epitope-encoding regions from Tp1**

The recombinant Tp1 plasmid DNA was cut with restriction enzyme Apa I to generate Exonuclease III resistant 3'-protruding termini. The linearized plasmid DNA was then purified with phenol-chloroform before digestion with the restriction enzyme Xho I that will cut at the single and unique vector cloning site Xho I located between Not I and Apa I sites (NB: the Tp1 cDNA insert does not contain neither Apa I nor Xho I sites). Digestion of the DNA with Xho I generated an exonuclease III sensitive recessed 3'-terminus (only at the 3'side of the cDNA insert). The linearized plasmid DNA was purified as described previously, 100 µg of DNA was digested at 37 °C with 1 unit of Exonuclease III in 20 µl reactions at varying times (from 0 to 30 minutes). Reactions were stopped by 10 minutes incubation at 75 °C, DNA was ethanol-precipitated, washed with 70% ethanol, air-dried and re-dissolved in 20 µl of 1x Mung Bean nuclease. Five units of Mung

Bean nuclease were added to each reaction, mixed and incubated at 30 °C for 1 hour to remove single-stranded extensions and create ligatable blunt ends. The DNA was again purified with phenol-chloroform, precipitated with ethanol, washed with 70% ethanol, air-dried and re-dissolved in 10µl of sterile distilled water. For re-ligation, 5µl of DNA was mixed with 1µl of 10x T4 ligation buffer, 3µl of water and 1µl of T4 DNA ligase, and incubated at 16 °C overnight.

Competent *E. coli* DH5α was transformed with 2 µl of the ligation and plated on agar plates as previously described. Plasmid DNA was isolated from single bacteria colonies as described and the inserts were excised by double digestion with BamH I and Bbs I (an isoschizomer of Bbv II). Plasmid clones with different cDNA inserts were selected for IFN-γ ELISpot screening.

Oligonucleotide primers were generated and used in PCR to clone small overlapping fragments of a 600 bp region containing the CTL epitope, into the mammalian expression plasmid vector pTargetT (Promega). Clones generated were then analysed by IFN-γ ELISpot.

### **1.12 Identification of Tp1 CTL epitopes with synthetic peptide libraries**

Peptide libraries (Cleaved PepSets; Mimotopes, Clayton, Australia) were generated for the 66 amino acid portion of Tp1 shown to contain the HD6 restricted CTL epitope. The PepSet libraries contained every 12mer, 11mer, 10mer and 9mer offset by 2 amino acids from the protein sequences. However, the peptides were prepared by truncations of the 12mers at the N-terminus and were supplied lyophilised with each tube containing a nominal 12mer and the 9,

10, 11mer truncations with the same C-terminus. Peptides were dissolved in 400µl 50% (v/v) DNA synthesis grade acetonitrile/water (Applied Biosystems, Warrington, UK). To aid the dissolution, tubes were held in a sonicator water bath for 2 x 10 min. Peptides were aliquoted into labelled cryopreservation tubes (Greiner) and stored at -20°C. For screening with CTL, peptides were prepared at 10µg/ml in complete RPMI-1640 and 10µl added to triplicate wells of an ELISpot plate, coated, washed and blocked as described above. Autologous iSF or P815 cells stably expressing the BoLA class I HD6 (P815-HD6) or JSP-1 (P815-JSP-1) were adjusted to a density of  $4 \times 10^5$ /ml and 50µl added to wells containing peptides. The plates were incubated at 37°C for 1 hour before CTL, prepared as described above for screening transfectants, were added 50µl/well. Plates were incubated for 20 hours at 37°C and then developed as described above. Based on the results of the screening with the Tp1 PepSet, individual 9, 10 and 11mer peptides were synthesised in order to define the CTL epitopes. Peptides were prepared and screened using the IFN- $\gamma$  ELISpot as described above.

Peptide-pulsed iSF and BoLA class I P815 transfectants were prepared as targets for  $^{51}$ Chromium release assays by incubating  $2 \times 10^6$  iSF or P815 cells overnight in T25 tissue culture flasks (Costar) with Tp1 peptides diluted to 1µg/ml in complete DMEM. Cells were harvested, labelled and assayed as described above.

### 1.13 Detection of Tp1 specific ex vivo CD8<sup>+</sup> T cell responses from immun cattle after challenge with *T. parva* sporozoites

Bull BV115, whose schizont specific CTL lines had been shown to recognise Tp1, was challenged with a lethal dose of *T. parva* (Muguga) sporozoites. Cryopreserved sporozoites (Stabilate # 4133) were thawed and diluted 1/20 as previously described. Animals were challenged by subcutaneous injection of 1ml of diluted sporozoites 2cm above the right parotid lymph node. Animals were monitored daily for changes in rectal temperature and from day 5 post challenge lymph node biopsies were taken using a 21G needle. Giemsa stained biopsy smears were examined for the presence of schizont infected cells and scored on a scale of 1-3. Animals were bled on day 0 and daily from day 6 to 13 and PBMC were isolated as described above. CD8<sup>+</sup> T cells and CD14<sup>+</sup> monocytes were purified from PBMC by MACS magnetic cell sorting according to the manufacturers instructions (Miltenyi Biotec, Gergisch Gladbach, Germany). CD8<sup>+</sup> T cells were sorted indirectly using a monoclonal antibody specific for bovine CD8 (IL-A105) followed by incubation with goat anti-mouse IgG microbeads (Miltenyi Biotec). CD14 monocytes were sorted directly by incubation with CD14 microbeads (Miltenyi Biotec). PBMC and CD8<sup>+</sup> T cells were added to wells ( $2.5 \times 10^5$ /well) of coated/blocked ELISpot plates and stimulated with autologous TpM ( $2.5 \times 10^4$ /well) or Tp1 peptides (1µg/ml final concentration). Purified monocytes were additionally added ( $2.5 \times 10^4$ /well) to wells containing peptide and CD8<sup>+</sup> T cells. ELISpot plates were incubated and developed as described above.

In order to recall Tp1 peptide specific CTL responses, PBMC were stimulated with autologous TpM 14 days post-challenge as described above. Viable cells were harvested 7 days post-stimulation and lytic activity against TpM and Tp1 peptide pulsed uninfected T cell blasts assessed as described above.

## **2 Results of screening for CTL target antigens**

### **2.1 Optimisation of IFN- $\gamma$ ELISpot for the recognition of target antigens by schizont specific CTL**

The ability of the IFN- $\gamma$  ELISpot to detect the recognition of TpM by CTL was first assessed using a CD8<sup>+</sup> polyclonal CTL line from animal F100. Fourteen days post-stimulation, CTL were added (5000/well) to coated/blocked ELISpot wells containing 25,000 irradiated autologous schizont and the formation of IFN- $\gamma$  spots assessed after a 20-hour incubation. CO-culture of CTL with TpM resulted in significant release of IFN- $\gamma$ . Pre-incubating the CTL for 30 min with a mAb against BoLA class I completely inhibited the IFN- $\gamma$  response whilst mAbs against MHC class II or the irrelevant CD21 antigen had no effect. Significantly, there was almost no spontaneous release of IFN- $\gamma$  from CTL cultured without TpM. This TpM line did not constitutively express TpM and no IFN- $\gamma$  spots could be attributed to the TpM.

In an attempt to replicate the transient transfection situation, where CTL would be co-cultured with COS-7 or iSF of which only a small proportion of cells would be

expressing the target antigen, TpM were titrated in COS-7 or iSF and co-cultured with CTL in IFN- $\gamma$  ELISpot plates. The stimulator population was fixed at an input of 40,000/well, with only the proportions of TpM and COS-7/SF varying. This cell input was adopted since it was thought to mirror the numbers of APC that would be co-cultured with CTL after transient transfection. Initially different CTL inputs were tried against TpM titrations with the aim of identifying the minimum CTL input required to detect significant responses to 1-3% TpM. A CTL input of 10,000/well was determined to be optimal since it could elicit significant responses to less than 1% TpM and it was practically feasible to raise such CTL numbers for screening experiments.

Further titration experiments were performed with CTL clones from F100 to confirm that with these CTL and APC inputs the IFN- $\gamma$  ELISpot was meeting or exceeding the desired sensitivity level. With all the clones tested the IFN- $\gamma$  ELISpot could still detect recognition of target cells when they constituted only 0.1% of the total cell population.

The ELISpot assay worked well with little background noise and met the sensitivity requirements when TpM were titrated in COS-7 cells but it was important to determine that the assay performed as well when the TpM were titrated in autologous SF. Neither primary nor iSF significantly affected the background levels or the sensitivity of the ELISpot assay.

In advance of the initiation of screening for CTL target antigens by the transient transfection of COS-7 cells and iSF, the efficiencies of COS-7 and iSF

transfection in 96 well TC plates was assessed using GFP as a reporter gene. Whilst there was considerable variation in transfection efficiencies between cell lines and between experiments, COS-7 consistently transfected better than iSF with efficiencies varying from 5-50% whereas for iSF transfection efficiency ranged between 0.5 – 20%. The transfection efficiency of iSF was assessed to be good enough to allow the presentation and identification of transfected schizont cDNA.

## **2.2 Evaluation of IFN- $\gamma$ Bioassay as a complementary read-out system for CTL recognition of target antigens**

Bovine vascular EC were stained for surface class II MHC expression following culture for 48 h in the presence of media containing recombinant IFN- $\gamma$ . The sensitivity to rBoIFN- $\gamma$  was determined and found to be between 100 and 10 pg/ml. Comparisons were made between IFN- $\gamma$  bioassay and ELISpot by co-culturing *T.parva*-specific CTL with fixed number of autologous iSF containing varying proportions of target autologous TpM and skin fibroblasts. There was a good correlation between the two assays. Both assays detected production of IFN- $\gamma$  in co-cultures containing as low as 1% TpM.

## **2.3 Identification of CTL target antigens**

### **2.3.1 Tp1**

#### **a) Identification**



A CD8<sup>+</sup> polyclonal T cell line from Friesian bull BV115 was generated and maintained using an HD6 expressing 4229 TpM (4229 TpM). This CTL line, termed BV115 (4229 TpM) CD8<sup>+</sup> polyclonal CTL line, lysed both autologous TpM and 4229 TpM showing that the line included schizont specific HD6 restricted CTL.

This CTL line was used to screen the 1000 pools of 50 schizont cDNA (B & C series) following co-transfection of COS-7 cells with schizont cDNA pools and a pcDNA3 construct expressing the full length cDNA for HD6. The CTL line responded to 10 cDNA pools and the results of screening pools B1 – B200 are shown in Figure 3. Six of the ten putative positive pools were subjected to resolution and by way of illustration the results of screening the 48 resolved pools from the putative positive pool B162 are shown in Figure 4. The 3-way matrix was decoded and the single cDNA were screened, the results from screening the single cDNA originating from pool B162 are shown in Figure 5. Of the five single cDNA only three were recognised by the CTL line, all five cDNA were sequenced and found that the positive cDNA were identical whilst the negative cDNA were unrelated. It is likely that the negative cDNA were contaminated with the positive cDNA during propagation of the *E. coli* in microtitre plates since they were always grown in wells adjacent to wells containing the positive cDNA.

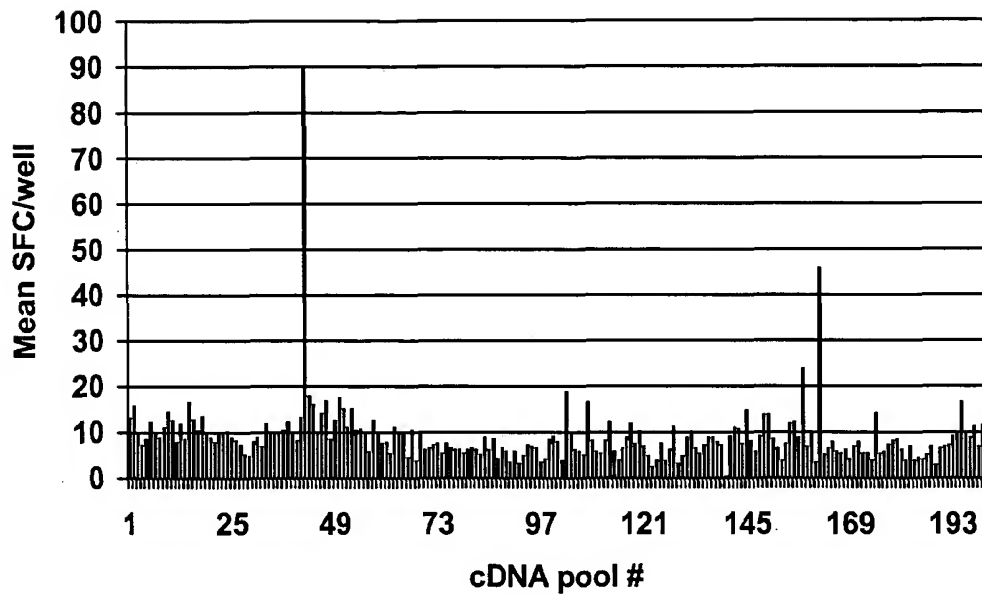


Figure 3: Screening of schizont cDNA pools (B series) with BV115 (4229 TpM) CD8 polyclonal CTL. COS-7 cells co-transfected with BoLA class I HD6 cDNA and schizont cDNA pools were cultured with CTL and recognition assessed by IFN- $\gamma$  ELISpot. Responses are presented as mean spot forming cell (SFC)/well. Of the 200 cDNA pools tested, pools B42 and B162 were selected as putative positives and subjected to resolution.

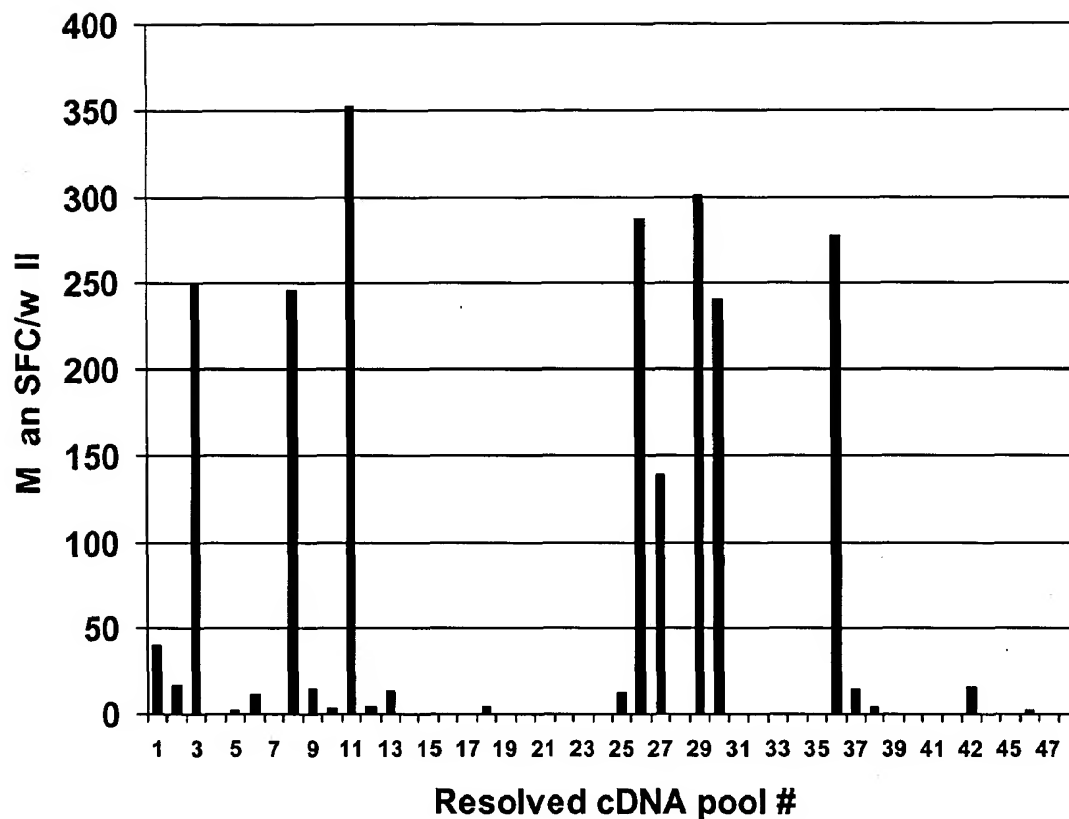


Figure 4: Screening of resolved schizont cDNA pool B162 with BV115 (4229 TpM) CD8 polyclonal CTL. 48 pools of 10 cDNA derived were constructed in a 3 way matrix and co-transfected with HD6 into COS-7 cells. Recognition of transfectants by CTL was assessed by IFN- $\gamma$  ELISpot. IFN- $\gamma$  production is presented as mean number of spot forming cell (SFC)/well. Of the 48 cDNA pools, 8 were positive, the 3 way matrix was then decoded to reveal 5 single cDNA.

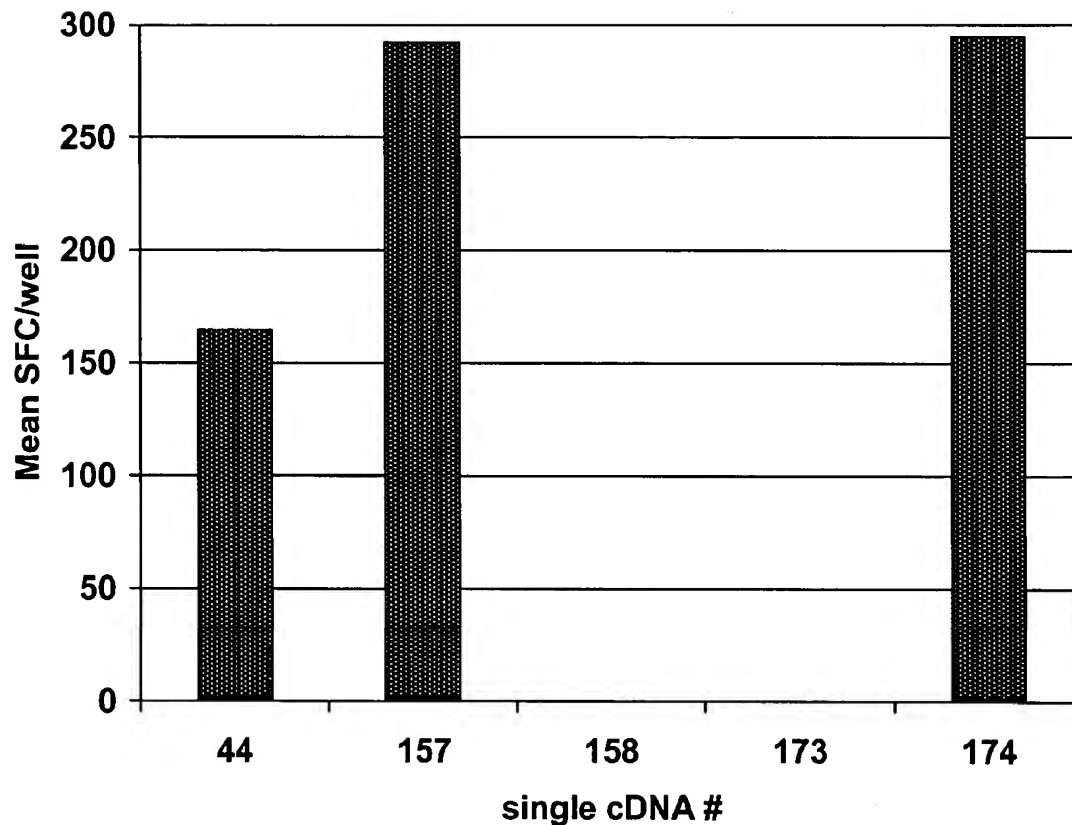


Figure 5: Screening of single schizont cDNA with BV115 (4229 TpM) CD8 polyclonal CTL following resolution of schizont cDNA pool B162. Five single cDNA decoded from the 3 way matrix were co-transfected with HD6 into COS-7 cells. Recognition of transfectants by CTL was assessed by IFN- $\gamma$  ELISpot. IFN- $\gamma$  production is presented as mean number of spot forming cells (SFC)/well. Of the 5 single cDNA, 3 were positive, the 3 had identical sequences and were different from the 2 negative cDNA.

The response to the other resolved cDNA pools were due to this same cDNA, which was termed Tp1. Specific PCR primers were designed and the remaining putative positive pools were found to be positive for Tp1 cDNA. HD6 restriction recognition of Tp1 was assessed by co-transfection of Tp1 with HD6 or another BoLA class I cDNA, KN104 or transfection of Tp1 alone, Figure 6 shows that only when COS-7 cells were co-transfected with HD6 and Tp1 did the CTL secrete IFN- $\gamma$ . The ability of autologous immortalised skin fibroblasts to present Tp1 was assessed by transfection and compared to co-transfected COS-7 cells. Figure 7 shows that immortalised skin fibroblasts presented Tp1 to the CTL although the response was less than induced by the COS-7 cells, probably due to the reduced transfection efficiency of skin fibroblasts. Recognition of Tp1 transfected COS-7 cells and immortalised skin fibroblasts was also confirmed by the use of the IFN- $\gamma$  bioassay (Figure 8).

The ability of BV115 (4229 TpM) CD8 polyclonal CTL line to lyse Tp1 transfected

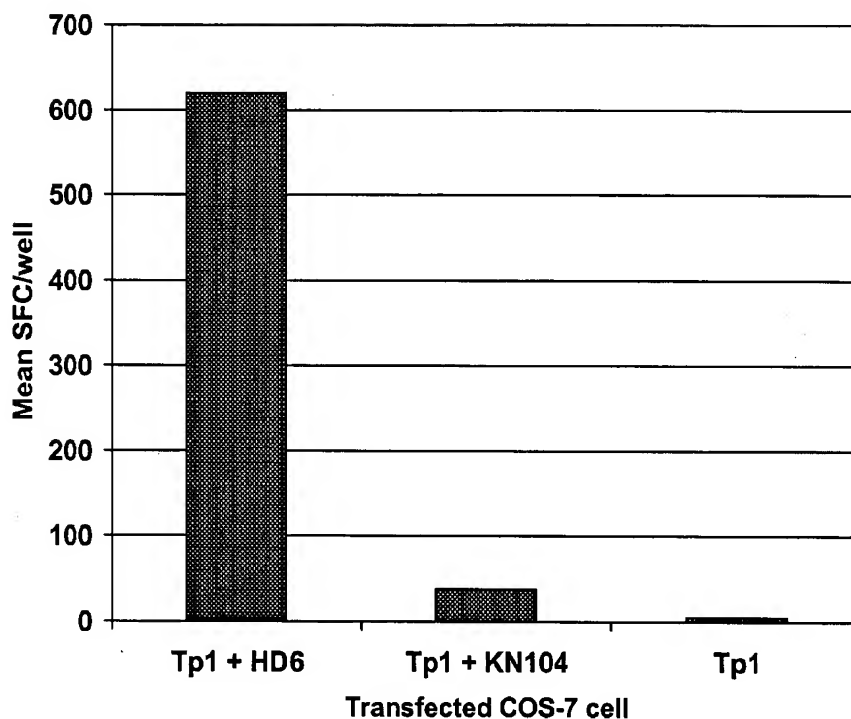


FIGURE 6

targets was then assessed using a  $^{51}$ Chromium release assay. COS-7 cells co-transfected with HD6 and Tp1, and Tp1 transfected autologous iSF were labelled and used as target cells. CTL efficiently lysed 4229 TpM and Tp1 transfected cells (Figure 9). The lysis of Tp1 expressing targets was inhibited by the presence of a blocking monoclonal antibody against BoLA class I. The CTL did not lyse cells transfected with the irrelevant *T. parva* antigen PIM.

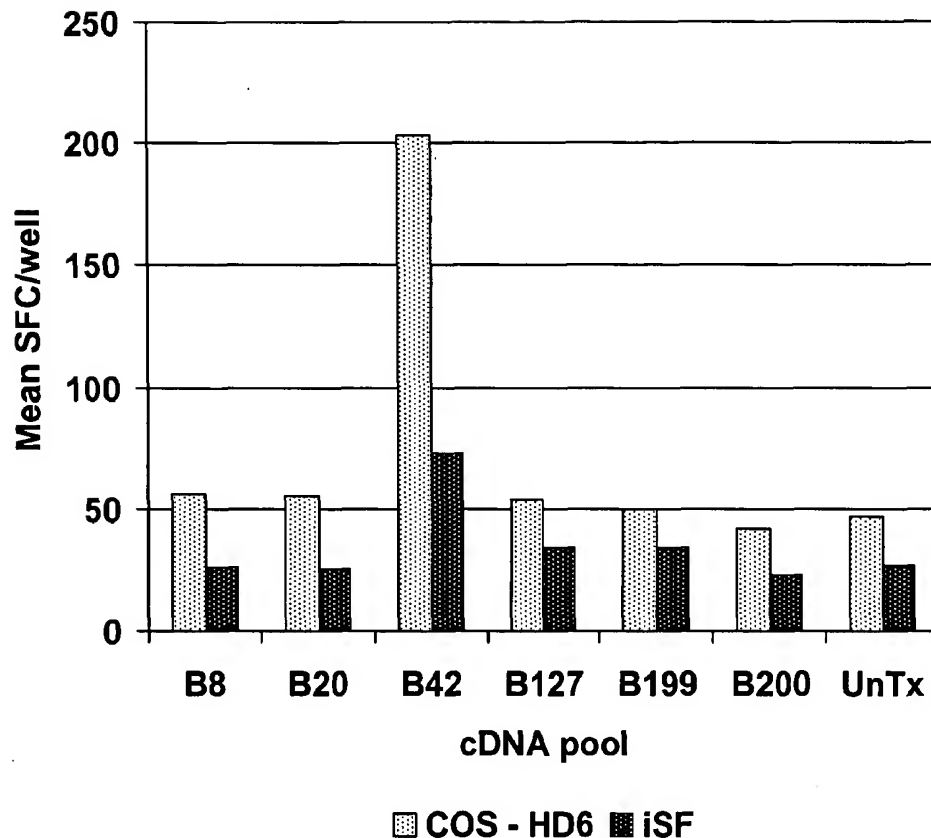


Figure 7: Comparison of ability COS-7 and iSF to present target antigen containing schizont cDNA pools to CTL. COS-7 were co-transfected with HD6 cDNA and schizont cDNA pools, whilst BV115 iSF were transfected with the schizont cDNA pools alone. Recognition of transfectants by BV115 (4229) CD8+ polyclonal CTL was assessed by IFN- $\gamma$  ELISpot. IFN- $\gamma$  production is presented as mean number of spot forming cells (SFC)/well. Of the 7 pools tested only one, pool B42, gave a response above that of the untransfected cells. This pool was known to contain the Tp1 cDNA. The weaker response to transfected iSF is likely due to an inferior transfection efficiency compared to COS-7.

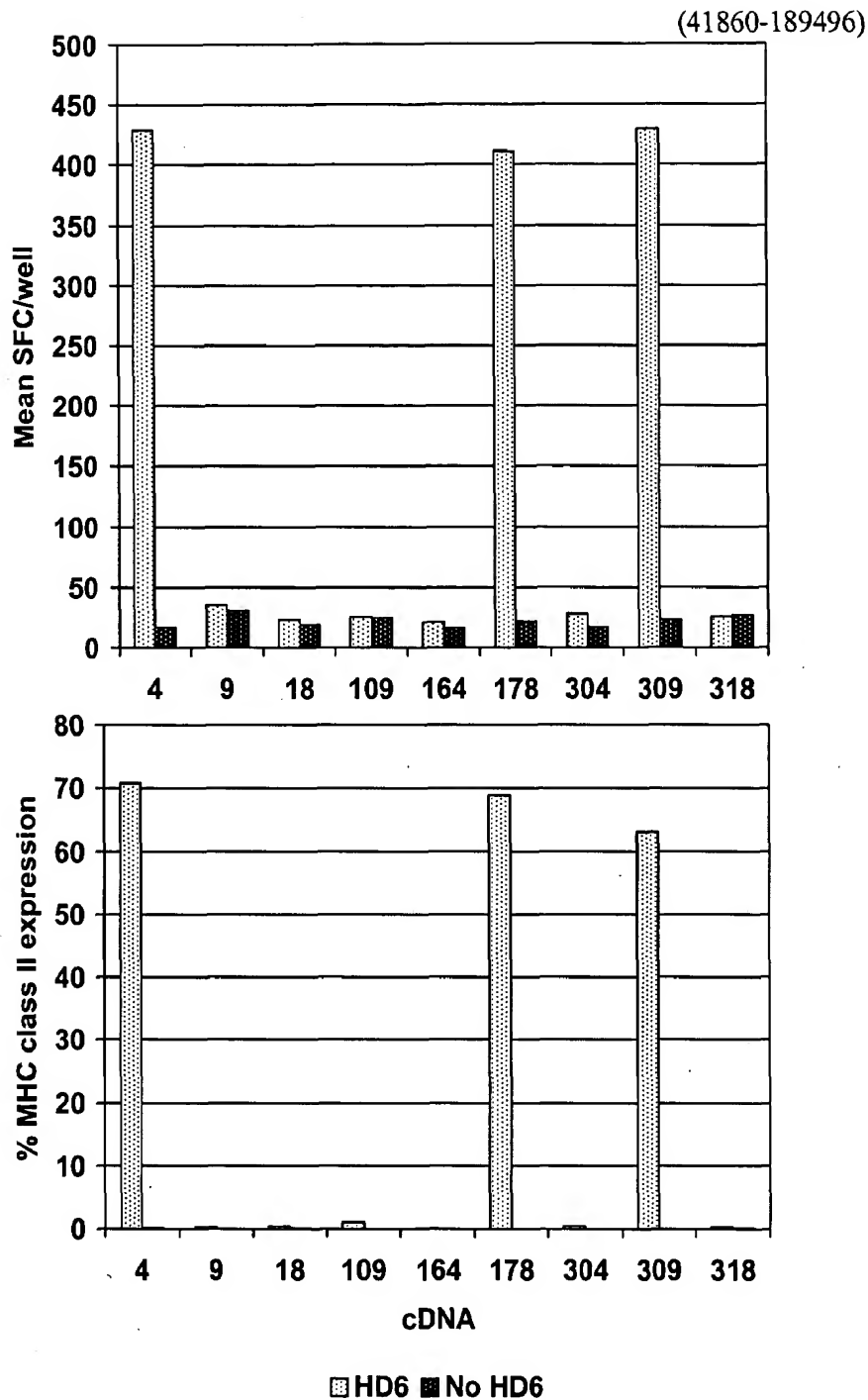


Figure 8: Comparison of detection of CTL IFN- $\gamma$  release by ELISpot (A) and Bioassay (B) following recognition of Tp1 transfected COS-7 cells. Single cDNA isolated following resolution of positive cDNA pool B42 were transfected into COS-7 cells with or without HD6 cDNA. Recognition of transfectants by BV115 (4229) CD8<sup>+</sup> polyclonal CTL was assessed by release of IFN- $\gamma$  using both an ELISpot and bioassay. IFN- $\gamma$  production as detected by ELISpot is presented as mean number of spot forming cells (SFC)/well and IFN- $\gamma$  bioactivity presented as % MHC class II expression. Of the 9 cDNA tested, both assays showed HD6 dependent recognition of 3 (#4, 178 and 309). All cDNA were sequenced and the positive cDNA identified as Tp1.



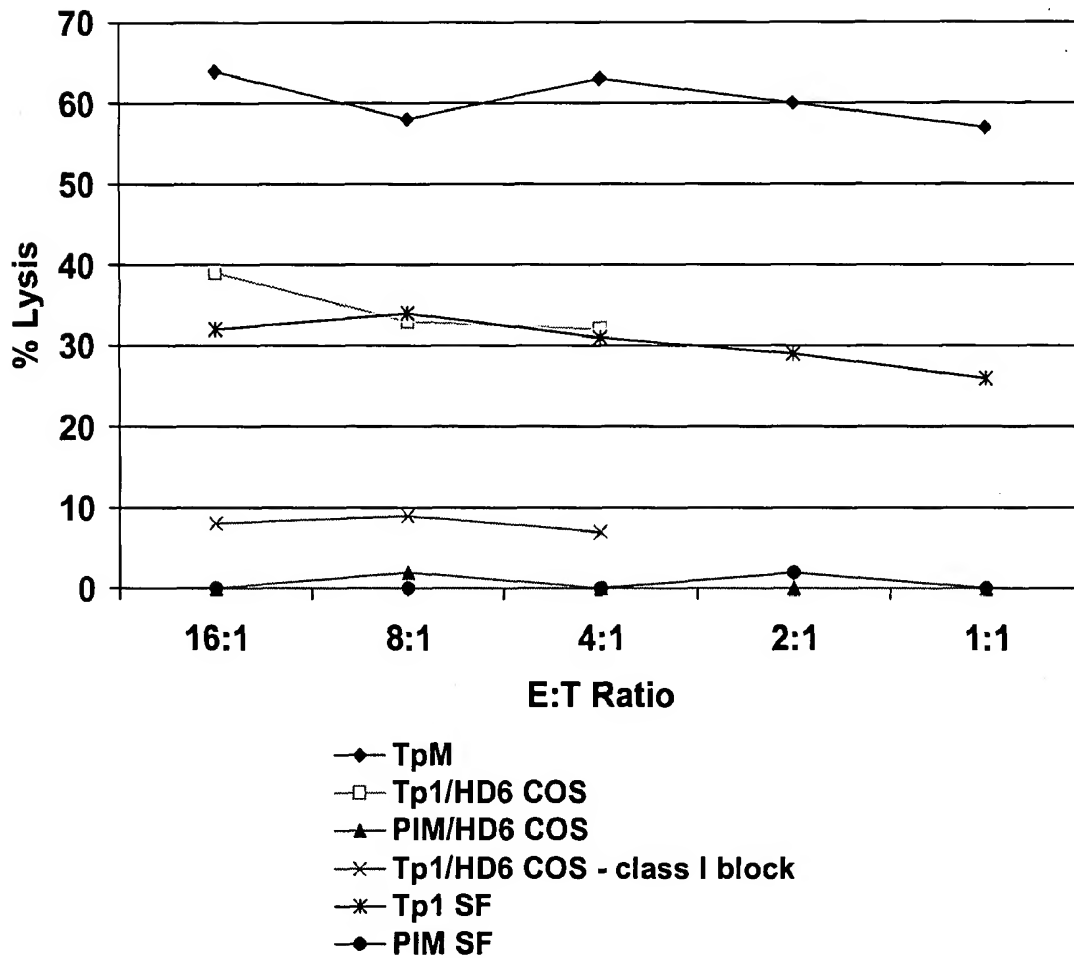


Figure 9: Lysis of Tp1 transfected COS-7 cells (co-transfection with BoLA class I HD6 cDNA) and immortalised autologous skin fibroblasts (SF) by the schizont specific BV115 (4229 TpM) CD8+ polyclonal CTL line. HD6 expressing 4229 TpM line was included as a positive control and PIM transfected COS-7 cells and iSF included as negative controls. MHC class I restriction was assessed by pre-incubating Tp1 transfected COS cells with an anti-BoLA class I mAb (MHC I block).

b) Mapping of the HD6 restricted Tp1 CTL epitope

Following Exonuclease III digestion, six plasmids containing Tp1 inserts of differing sizes (Tp1 Del1 to Del6) were co-transfected with HD6 into COS-7 cells and recognition of plasmid clones by BV115 (4229 TpM) CD8 polyclonal CTL line assessed by IFN- $\gamma$  ELISpot. Positive ELISpot responses were observed against clones Tp1 Del1 and Del2. Both ELISpot positive and negative plasmid clones were sequenced and compared. Figure 10 shows a comparison of the predicted amino acid sequence of Tp1 with the deleted clones. The sequence comparison located the HD6 restricted CTL epitope within a 101 amino acid fragment of Tp1. Oligonucleotide primers were generated and used in PCR to clone a 600bp fragment of Tp1, overlapping the epitope-encoding region (Tp1.6). IFN- $\gamma$  ELISpot confirmed CTL recognition of this portion of Tp1 and so small overlapping fragments (Tp1.1-Tp1.5) of the 600 bp region were cloned into the mammalian expression plasmid vector pTargetT. Clones were then analysed by ELISpot and the CTL epitope was further narrowed down to a 200 bp DNA fragment corresponding to 66 amino acids (Figure 11).

	10	20	30	40	50	(41860-189496)
80						
Tp1 ORF	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
Tp1 Del1	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
Tp1 Del2	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
Tp1 Del3	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
Tp1 Del4	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
Tp1 Del5	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
Tp1 Del6	MRVKKVLLYT	LPVVGILLAG	SLIIFNFVRK	RPEKEEELKP	PSALEDELKK	REEESRKRME
	90	100	110	120	130	140
Tp1 ORF	LEKREKEVVD	EFAKHLKKPE	ERLPKIILTL	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
Tp1 Del1	LEKREKEVVD	EFAKHLKKPE	ERLPKIILTL	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
Tp1 Del2	LEKREKEVVD	EFAKHLKKPE	ERLPKIILTL	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
Tp1 Del3	LEKREKEVVD	EFAKHLKKPE	ERLPKII...	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
Tp1 Del4	LEKREKEVVD	EFAKHLKKPE	ERLPKII...	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
Tp1 Del5	LEKREKEVVD	EFAKHLKKPE	ERL.....	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
Tp1 Del6	LEKREKEVVD	EFAKHLKKPE	ERL.....	DSGFPTVDPI	TYTSGVYMVA	VSKTTFTSDS
	170	180	190	200	210	220
Tp1 ORF	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
Tp1 Del1	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
Tp1 Del2	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
Tp1 Del3	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
Tp1 Del4	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
Tp1 Del5	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
Tp1 Del6	FGGKTYTIKP	IEATMATSIA	FAADPGFCYF	LLIPGPDSPK	IFFKNDGDKF	LRCVGYPKVK
	250	260	270	280	290	300
Tp1 ORF	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
Tp1 Del1	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
Tp1 Del2	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
Tp1 Del3	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
Tp1 Del4	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
Tp1 Del5	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
Tp1 Del6	APPGVKPEAP	TPTPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN	KFMIGSQEVK
	330	340	350	360	370	380
Tp1 ORF	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
Tp1 Del1	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
Tp1 Del2	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
Tp1 Del3	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
Tp1 Del4	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
Tp1 Del5	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
Tp1 Del6	GSFTPTSFSI	GDLPQTGLPV	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIFY	KNQSRKYLNK
	410	420	430	440	450	460
Tp1 ORF	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
Tp1 Del1	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
Tp1 Del2	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
Tp1 Del3	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
Tp1 Del4	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
Tp1 Del5	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
Tp1 Del6	DSIFNDFVTV	NLSRRSDYYR	NGTGTSEIEQ	TLDNMVYVEP	DTPCAGWTTY	IHKLEEGGEG
	490	500	510	520	530	540
Tp1 ORF	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL
Tp1 Del1	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL
Tp1 Del2	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL
Tp1 Del3	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL
Tp1 Del4	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL
Tp1 Del5	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL
Tp1 Del6	FPMGKVSIVN	VYGKNDPLS	YAPSIFSIVR	EDGIQIFYVR	AYSQYLLDSS	VNPQNLPQKL

Figure 10: Deduced amino acid sequences of Exonuclease III deleted Tp1 clones. Clones were co-transfected with HD6 cDNA and recognition by CTL determined by IFN- $\gamma$  ELISpot. Only the Tp1 orf and clones Tp1 Del1 and Del2 were recognised implying that the HD6 restricted Tp1 epitope lay between amino acid position 141 and 241.

	200	210	220	230	240
Tp1.1	<b>m</b> PGPDSKP	IFFKNDGDKF	LRCVGYPKVK	EEMLEMATKF	NRLPKGVEIP
Tp1.2					
Tp1.3					
Tp1.4	<b>m</b> PGPDSKP	IFFKNDGDKF	LRCVGYPKVK	EEMLEMATKF	NRLPKGVEIP
Tp1.5					
Tp1.6	<b>m</b> PGPDSKP	IFFKNDGDKF	LRCVGYPKVK	EEMLEMATKF	NRLPKGVEIP
	250	260	270	280	290
Tp1.1	APPGVKPEAP	TPTPTTIT			
Tp1.2	<b>m</b> P	TPTTITPS	VPPTIPTPIT	PSAPPTTPPT	GLNFNLTQVN
Tp1.3					
Tp1.4	APPGVKPEAP	TPTTITPSVP	PTIPTPITPS	APPTTPPTGL	NFNLTQVNKF
Tp1.5	<b>m</b> P	TPTTITPSVP	PTIPTPITPS	APPTTPPTGL	NFNLTQVNKF
Tp1.6	APPGVKPEAP	TPTTITPSVP	PTIPTPITPS	APPTTPPTGL	NFNLTQVNKF
	300	310	320	330	340
Tp1.1					
Tp1.2	KFMIGSQEVK	LNITHEYEGV	YEAHKYFI		
Tp1.3		<b>m</b> GV	YEAHKYFIER	GSFTPTSFSI	GDLPQTGLPV
Tp1.4	KFMIGSQEVK	LNITHEYEGV	YEAHKYFIER	GSFTPTSFSI	GDLPQTGLPV
Tp1.5	KFMIGSQEVK	LNITHEYEGV	YEAHKYFIER	GSFTPTSFSI	GDLPQTGLPV
Tp1.6	KFMIGSQEVK	LNITHEYEGV	YEAHKYFIER	GSFTPTSFSI	GDLPQTGLPV
	350	360	369		
Tp1.1					
Tp1.2					
Tp1.3	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIF		
Tp1.4					
Tp1.5	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIF		
Tp1.6	NQTVDTIVVY	FHRVTMGEPV	GIPLIVLIF		

Figure 11: Deduced amino acid sequences of deleted Tp1 clones. The small m in bold represents the methionine added by the artificial ATG start codon by PCR. Constructs were co-transfected with HD6 cDNA and recognition by CTL determined by IFN- $\gamma$  ELISpot. Clones Tp1.1, 1.4 and 1.6 were recognised by CTL thus narrowing down the epitope containing region to 66 amino acids.

A cleaved PepSet library of 28 overlapping peptides each were synthesised to encompass the 66 amino acids encoded by the Tp1.2 insert (Mimotopes). Recognition of Tp1 peptides was assessed by IFN- $\gamma$  ELISpot, using autologous immortalised skin fibroblasts or murine mastocytoma P815 cells (ATCC #TIB-64, ATCC, Manassas, VA USA) stably expressing HD6 as antigen-presenting cells. Figure 12 shows the results of screening the cleaved Tp1 Pepset with BV115 (4229 TpM) CD8 polyclonal CTL line and BV115 CTL clones. The recognition of peptides #10 and #11 suggested that the epitope fell within the region RCVGYPKVKEEMLE. All possible 9, 10 and 11mers were then synthesised for this sequence and screened against the polyclonal line and clones. Figure 13 shows the responses of two BV115 CTL clones to individual Tp1 peptides. Both clones responded to peptide #24, the 11mer VGYPKVKEEML, suggesting that this was the minimal length HD6 restricted epitope of Tp1. This data was supported by <sup>51</sup>Chromium release assays that demonstrated significant CTL lysis of Tp1 peptide #24 pulsed iSF or P815-HD6.

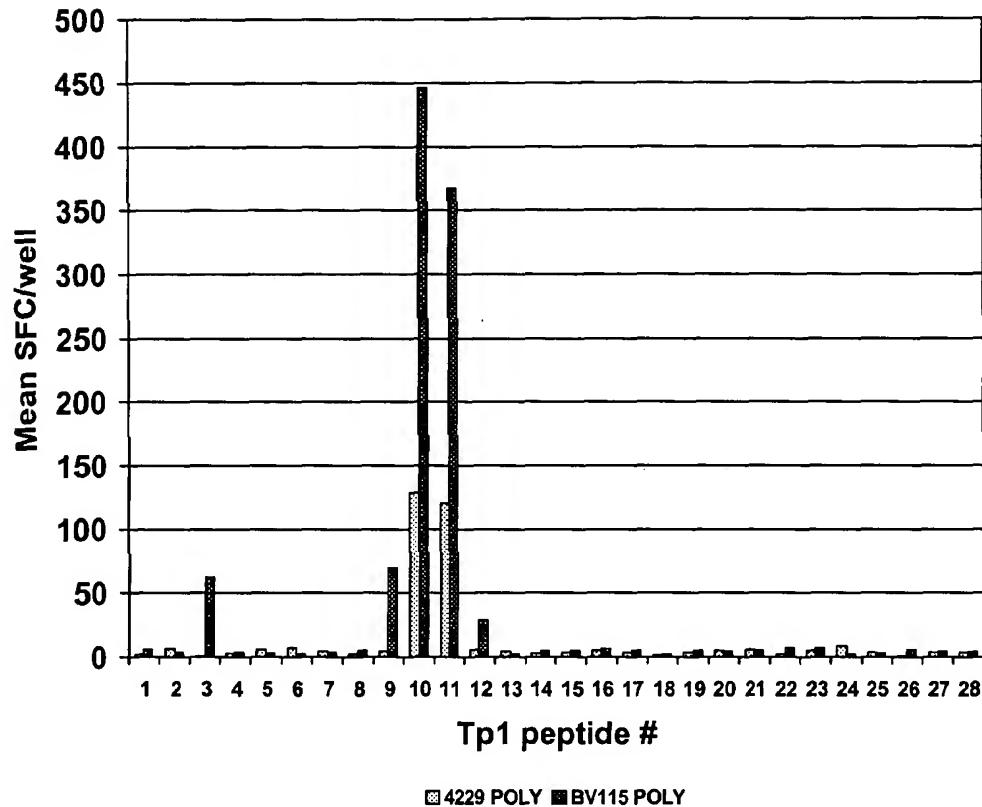


Figure 12: Fine mapping of the HD6 restricted Tp1 CTL epitope using synthetic peptide library. Twenty eight 12mer peptides overlapping by two amino acids encompassing the 66 amino acids encoded by the Tp1.2 insert were synthesised and used at a final concentration of 1µg/ml to pulse BV115 iSF. Recognition of peptide pulsed iSF by two polyclonal CTL lines derived from BV115, one maintained on 4229 TpM (4229 poly) the other on autologous BV115 TpM (BV115 poly), was assessed by IFN-γ ELISpot. Significant responses were observed against peptides #10 and 11 that corresponded to the amino acid sequence RCVGYPKVKEEMLE.

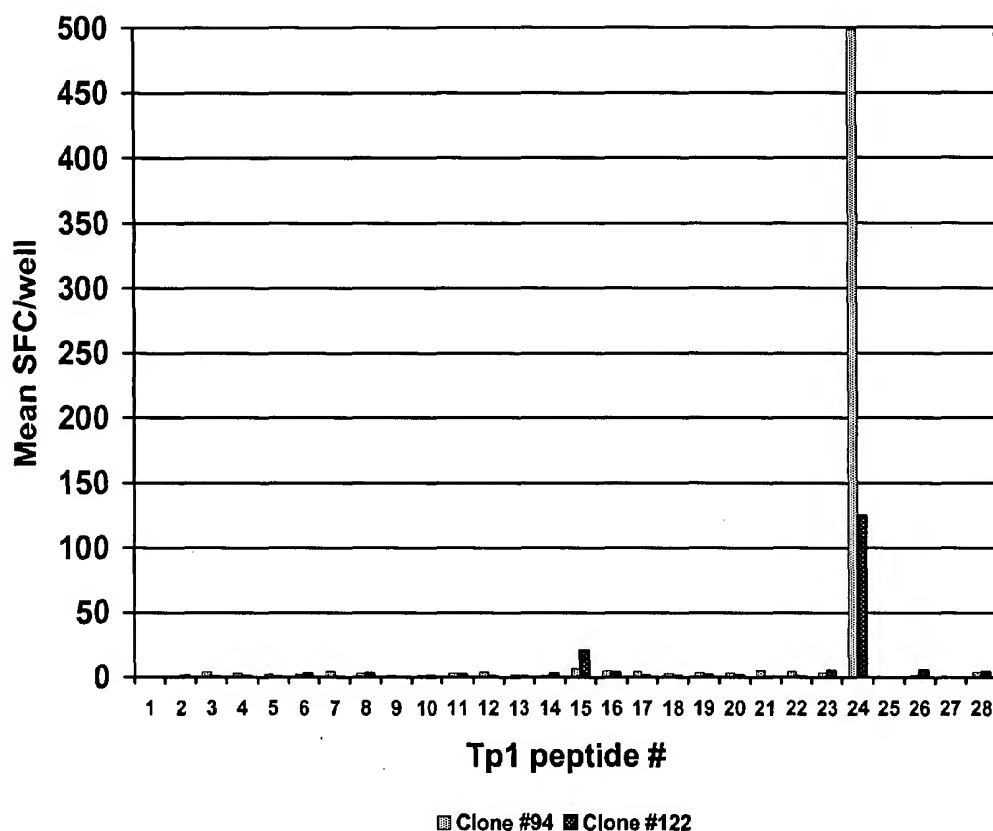


Figure 13: Identification of the HD6 restricted Tp1 CTL epitope using synthetic peptides. Twenty seven peptides comprising of all the possible 9, 10 and 11mers from the sequence FLVGYPKVKEEMLEMA were synthesised and used at a final concentration of 100pg/ml to pulse P815 cells stably expressing HD6 (P815/HD6). Recognition of peptide pulsed P815/HD6 by two HD6 restricted BV115 CTL clones, #94 and #122, was assessed by IFN- $\gamma$  ELISpot. Significant responses were only observed against peptide #24, suggesting that the HD6 was the 11mer VGYPKVKEEML.

Tp1 was cloned from the Marikebuni stock of *T. parva*. The deduced amino acid sequence was compared to that of Tp1 from the Muguga stock and found to be 95.8% identical (Figure 14). The ability of CTL to recognise the Marikabuni and Muguga Tp1 was assessed using the ELISpot. BV115 (4229 TpM) CD8 polyclonal CTL line only recognised the Muguga Tp1. This is explained by a double amino acid substitution at the C-terminus of the predicted HD6 restricted epitope.



	10	20	30	40	50	60
Tp1Muguga	MRVKVLLYTLPPVVGILLAGSLIIFNFVRKRPEKEEELKPPSALEDELKKREESRKRME					
Tp1Muguga	.....					
Tp1Muguga	10	20	30	40	50	60
Tp1Muguga	70	80	90	100	110	120
Tp1Muguga	EMQKEILEKKLREGKKALEELEKREKEVVDEFKHLKKPEERLPKIILTLDSGFPTVDPI					
Tp1Muguga	.....					
Tp1Muguga	70	80	90	100	110	120
Tp1Muguga	130	140	150	160	170	180
Tp1Muguga	TYTSGVYMAVSKTTFTSDSDLVDFHTLLGIKFLVTGVQFGGKTYTIKPIEATMATSIA					
Tp1Muguga	.....					
Tp1Muguga	130	140	150	160	170	180
Tp1Muguga	190	200	210	220	230	240
Tp1Muguga	FAADPGFCYFLLIPGPDSPKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIP					
Tp1Muguga	.....					
Tp1Muguga	190	200	210	220	230	240
Tp1Muguga	250	260	270	280	290	300
Tp1Muguga	APPGVKPEAPTPTPTTITPSVPPTIPTPITPSAPPTTPTGLNFNLTQNKFMIGSQEVK					
Tp1Muguga	.....					
Tp1Muguga	250	260	270	280	290	300
Tp1Muguga	310	320	330	340	350	360
Tp1Muguga	LNITHEYEGVYEAHKYFIERGSFTPTSFSIGDLPQTGLPVNQTVDTIVVYFHRVTMGEPV					
Tp1Muguga	.....					
Tp1Muguga	310	320	330	340	350	360
Tp1Muguga	370	380	390	400	410	420
Tp1Muguga	GIPLIVLIFYKNQSRKYLNGNGNWEESKALLFREELDYLDSIFNDFVTNLSRRSDYYR					
Tp1Muguga	.....					
Tp1Muguga	370	380	390	400	410	420
Tp1Muguga	430	440	450	460	470	480
Tp1Muguga	NGTGTSEIEQTLDMNVYVEPDTPCAGWTYYIHKLEEGGEGGIEKPFQIRQLWFSKQKFDI					
Tp1Muguga	.....					
Tp1Muguga	430	440	450	460	470	480
Tp1Muguga	490	500	510	520	530	540
Tp1Muguga	FPMGKVSIVNVYGNDEPLSYAPSIFSVIREDGIQIFYVRAYSQYLLDSSVNPQNLQPQL					
Tp1Muguga	.....					
Tp1Muguga	490	500	510	520	530	540
Tp1Muguga	NTL					
Tp1Muguga	..					
Tp1Muguga	TAE					

Figure 14: Comparison of deduced amino acid sequences of Tp1 from *T. parva* Muguga and Marikebuni stocks. There was 95.8% identity in 542 aa overlap. Significantly there were two amino acid differences at positions 226 and 227, the C terminal end of the 11 amino acid HD6 restricted epitope.

c) Evidence that Tp1 is recognised by CD8<sup>+</sup> T cells from an immune resolving a challenge infection

Bull BV115, whose CTL had recognised Tp1, was challenged with a lethal dose of *T. parva* (Muguga) sporozoites and the response of purified CD8<sup>+</sup> T cells to the HD6 restricted Tp1 epitope or control Tp1 peptide which did not contain any previously identified epitopes were measured longitudinally. The animal was solidly resistant to challenge with no fever or detectable parasitosis. From day 8 post-challenge, CD8<sup>+</sup> T cells responded specifically to the Tp1 epitope and were sustained over the period of observation (Fig 15). The kinetics of this response is comparable to that previously described for schizont-specific CTL precursors in blood following challenge of immune cattle with *T. parva* sporozoites (McKeever *et al* 1994). Attempts were made to detect Tp1 and TpM specific lytic responses directly in peripheral blood post-challenge but these failed. An experiment was instigated to first expand schizont specific CTL numbers by a single *in vitro* stimulation with TpM and then to assess Tp1 specific lysis. Stimulated cells exhibited high cytotoxic activity against both TpM and T cell blasts pulsed with the Tp1 epitope. These data suggest that a component of the protective CD8<sup>+</sup> T cell response is Tp1-specific highlighting the potential of this antigen as a vaccine candidate.

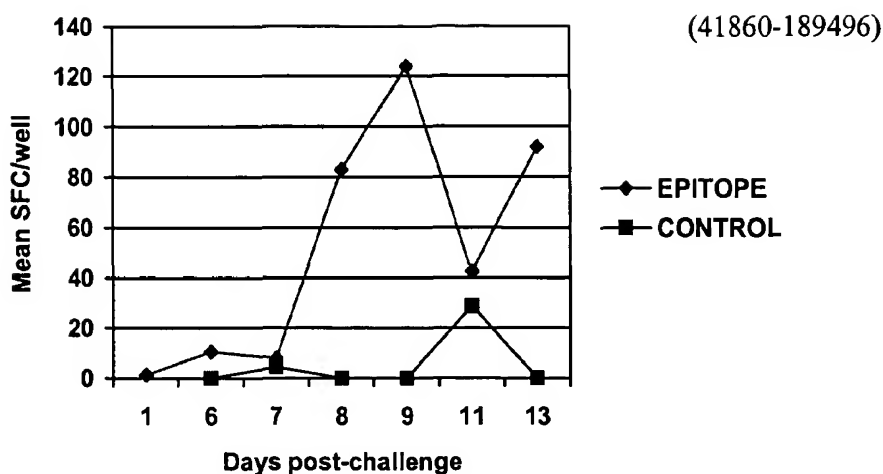


Figure 15: Tp1 specific CD8<sup>+</sup> T cell responses of an immune bull following challenge with *T. parva* sporozoites. CD8<sup>+</sup> T cells were purified from peripheral blood at various time-points and incubated with the HD6 restricted Tp1 11mer synthetic epitope (epitope) or a Tp1 peptide which had not previously been shown to contain a CTL epitope (control) and responses measured by IFN- $\gamma$  ELISpot assay.

### 2.3.2 Identification of other CTL target antigens

Four other CTL target antigens, namely Tp4, Tp5, Tp7 and Tp8, have been identified through screening and resolution of the schizont cDNA library using either the ELISpot assay or bioassay. These assays utilised CTLs generated from five cattle representing four class I MHC genotypes. Cells transfected with these four antigens were recognised and lysed specifically by CTLs. Further assays have been performed to define the CTL epitope in Tp5. Experiments have been carried out which demonstrate that Tp5 is recognised by *ex vivo* CD8<sup>+</sup> T cells obtained from an immune animal (BV050). These results are summarised in Table 4.

Antigen	CTL	Assay for ID	Lysis	Epitope	Ex vivo response
Tp4	BV057 BX063 BX065	ELISpot Bioassay	✓	Not done	Not done
Tp5	BV050	ELISpot	✓	✓	✓
Tp7	BW012	ELISpot	✓	Not done	Not done
Tp8	BX063 BX065	ELISpot	✓	Not done	Not done

Table 4: A summary of results showing the identification of further CTL target antigens.

## 2.4 Nucleotide and deduced amino acid sequences and putative identity of the 5 candidate antigens

Following confirmation of the candidates as CTL target antigens, the individual cDNA were sequenced and confirmed to be of *T. parva* origin by interrogating the *T. parva* genome sequence database. BLAST searches (Altschul *et al* 1990) of DNA and protein databases were performed and homologues of some of the antigens identified. SignalP (Nielsen *et al* 1997) and Tmpred ([http://www.ch.embnet.org/software/TMPRED\\_form.html](http://www.ch.embnet.org/software/TMPRED_form.html)) analyses were conducted to predict the presence of a signal peptide and transmembrane domain. Figures 16-20 show the DNA and deduced protein sequences of the five candidate antigens.

Nucleotide sequence of Tp1 ORF:

*aaggtaaagtatagatttttgacttttatttacttacctgtttgtataaaaattataagggtataaaatttaaccaaataa  
ccgatttaattgtaataatttaaagttttgtacagtataTGAGGGTCAAAAAAGTTTATTATATACACTTCCG  
GTTGTCGGTATTTTACTGGCTGGATCTTTGATTATATTTAATTTTCGTTAGGAAAAGACCGGAA  
AAAGAAGAGGAACCTCAAACCTCCTTCTGCATTAGAAGATGAACCTAAAAAACGTGAAGAAGA  
AAGCCGAAAACGCATGGAAGAAATGCAAAAGGAAATTCTCGAAAAAAGTTAAGAGAAGGTA  
AAAAAGCCTTGGAAGAACTTGAAAAACGTGAAAAAGAAGTGGTAGATGAGTTTGCAAAACAC*

CTCAAAAAACCTGAAGAAAGACTTCCTAAAATTATTCTTACATTGGATTCCGGTTTTCCAACA  
 GTTGATCCTATTACATATACTTCAGGAGTTTATATGGTAGCAGTTAGTAAAAACAACTTTTACCT  
 CAGATTACAGATCTTGTTGATTTTACTCACACACTGCTGGGCATAAAGTTTCTAGTTACTGGTG  
 TACAATTTGGTGGGAAAACATACACAATTAACCGATTGAAGCTACTATGGCCACTTCAATTG  
 CATTTGCCGCTGATCCTGGATTCTGTTATTTTCTATTAATACCAGGCCCTGACTCGAAACCA  
 ATATTCTTCAAAAACGACGGAGATAAATTTTTACGTTGCGTAGGGTATCCAAAGGTTAAA  
 GAAGAAATGCTAGAAATGGCTACAAAATCAATAGACTACCAAAGGGCGTGGAATACCT  
 GCACCTCCAGGAGTAAAACCAGAGGCTCCCACACCTACACCAACGACAATAACTCCTTCT  
 GTACCTCCTACTATACCAACGCCAATAACTCCTTCGGCACCTCCTACTACACCACCTACG  
 GGAATAATTTTAACTTGACAGTTCAGAACAAATTCATGATAGGTTTCGCAAGAAGTTAAG  
 TTAATATAACTCACGAATACGAGGGTGTATACGAAGCTCATAAATATTTTATTGAAAGG  
 GGCAGCTTTACCCCTACCTCATTCTCAATCGGTGATCTTCCACAAACAGGTCTTCCAGTA  
 AATCAACAGTGGATACAATTGTGGTATATTTCCATCGAGTAACGATGGGTGAACCCGTT  
 GGTATACCTCTAATTGTGTTAATCTTTTATAAAAACCAATCTAGAAAATATTTAAATAAG  
 GGAAATGGAACTGGGAAGAATCTAAAGCTCTATTATTTTCGTGAGGAACCTTGATTACTTA  
 GATTCCATTTTTAATGATTTTGTGACAGTAAACCTTTCTAGACGTTCTGATTATTATCGT  
 AATGGAACCTGGCACATCAGAAATTGAGCAAACGTTAGATATGAATGTTTACGTTGAACCT  
 GATACACCCTGTGCTGGATGGACAACGTATATACATAAATTAGAAGAAGGAGGTGAAGGA  
 GGAATTGAAAAACCTTTTCAAATTAGACAATTATGGTTTAGTAAACAGAAATTTGATATA  
 TTCCCAATGGGAAAAGTTTCAATCGTTAATGTTTATGGGAAAAACGATGAACCGCTATCT  
 TACGCTCCCTCAATTTTCAAGTGAATACGCGAAGATGGAATTCAAATATTTTATGTTTCT  
 GCTTACTCACAGTACTTGCTTGATTCAAGTGTTAATCCCCAAAATTTACCTCAAAAATTA  
 AACACCCTTTAG*atTTTTTaaaaaatcatgtaataataattgttttgaaaaaaaaaaaaaaaa*

#### Amino acid sequence of Tp1:

MRVKKVLLYTLPVVGILLAGSLIIFNFVRKRPEKEEELKPPSALEDELKKREEESR  
 KRMEEMQKEILEKKLREGKKALEELEKREKEVVDEFKHLKKPEERLPKIILTLD  
 SGFPTVDPITYTSGVYMVAVSKTTFTSDSDLVDFTHTLLGIKFLVTGVQFGGKTY  
 TIKPIEATMATSIAFAADPGFCYFLLIPGPDSPKPIFFKNDGDKFLRCVGYPKVKEE  
 MLEMATKFNRLPKGVEIPAPPGVKPEAPTPTPTTITPSVPPTIPTPITPSAPPTTP  
 PTGLNFNLTVQNKFMIGSQEVKLNITHEYEGVYEAHKYFIERGSFTPTSF SIGDL  
 PQTGLPVNQTVDTIVVYFHRVTMGEPVGIPLIVLIFYKNQSRKYLKNGNGNWE  
 SKALLFREELDYLDISFNDFVTNLSRRSDYYRNGTGTSEIEQTLDMNVYVEPD  
 TPCAGWTTYIHKLEEGGEGGIEKPFQIRQLWFSKQKFDIFPMGKVSIVNVYGKN  
 DEPLSYAPSIFSVIREDGIQIFYVRAYSQYLLDSSVNPQNLPQKLNTL

Figure 16: The nucleotide sequence of Tp1 cDNA. Tp1 cDNA contains a complete open reading frame (ORF) of 1632 nucleotides long shown in capital letters. The ATG start codon is at position 121 whereas the TAG stop codon is at position 1750. This ORF codes for a protein of 543 amino acids (aa). The 5' and 3' flanked untranslated regions (UTR) are in italics.

## Nucleotide sequence of Tp4 ORF:

ATGAGTCATTTAATGAACCTACCAATCCTTGTATTGAAGGAAGGCACTGATA  
 CATCCCAAGGCCAAGCTCAAATCATTAGTAATATCAACGCCTGTCAAGCTAT  
 TGTCGATTGTGTTAAACTACTCTAGGTCCTAGAGGAATGGACAAGTTGATA  
 CATACGGAGAGAGATGTGACGATAACCAATGACGGTGCTACTGTTTTGAAAT  
 TACTTGATATTACTCATCCTGCCGCTTCTGTTCTTGTTGATATCGCTAAATCA  
 CAAGATGATGAAGTCGGTGATGGGACTACTTCCGTTACTGTTCTAGCAGGT  
 GAGTTATTGAATGAAGCTAAGGCGTTTATATTGGATGGGATAAGTCCTCAGG  
 TTATCATAAAATACTATCGTGAAGCCTGTCAAGTTGCTTTAAATCTCATTGAT  
 AAAGTTGCCATTTCATCTCTCCAACAAATCCTCAACTGATAAGAAAGAACTACT  
 GATAAAATGTGCTGAAACTACTTTTAATTCAAAGTTATTGTCTGGTTATAAAA  
 CCTTTTTTGCCAAGATGGTTGTGGAGGCAGTGGCTACTTTGGATGAGGACTT  
 GGATGAGGATATGATTGGTGTAAAAAAGTCACTGGTGGTTCCTGTGAGGA  
 CTCCTCCTAGTCAAGGGTGTAGCATTCAAGAAAACCTTTCAGCTACGCTGG  
 GGCTGAACAACAGCCAAAGAAATTCGTCAATCCAAAGATTTTATTACTTAATT  
 TGGAATTGGAACCTCAAATCCGAAAAAGAAAACGCAGAAATTGTTATCAATAA  
 TCCACAAGAATATCAGAAGATAATAGATGCCGAGTATAGGATAATATTTGAG  
 AAGCTTGAGAATGCAGTGAAACTCGGTGCTAATGTAGTTTTATCTAAATTGC  
 CAATTGGTGATTTAGCAACACAATACTTTGCAGATAAAAATGTATTTTGTGCC  
 GGCCGGGTTGATGAAATGATCTTATAAGAACGAGTAAAGCTACTGGTGCTT  
 CTATTCAAACCACTCTCAATAACCTTTCAGTTGACGTCTTAGGAACTTGTGGT  
 GTGTTTGAGGAAGTGCAAATTGGGTCTGAACGTTACAATATGTTACAGATT  
 GCAAGAGTGCAAAAACCTGTACAATTGTGTTGAGAGGTGGAGGTCAGCAGT  
 TCATTGATGAATCTGAACGTTCACTCCATGACGCGATTATGATTGTCAGAAG  
 AGCAACTAAATGTAATACTATCCTTCCCGGAGCTGGTGCCATTGAGATGTTG  
 CTCTCAACTTATCTCCTCCACTATTCTCTCAACACTATTAATCCACAGACTC  
 TGTC AACCATGTAACTGCGTTAACTCCGTAAATCATGTTAATGGAGTTACTG  
 GGGTGAATAAGAGTCTGGTGGGTAAGAGGCACATAATAATGAACGGGTTTG  
 CAAAGGCATTGGAGTGTATCCAAGGAATTTAGCCACTAATTCTGGCTACAA  
 TTCAAATGATTTATTATCGATACTAAGAAATAAATACAATCAATTGGAAATAGT  
 CAATGGAGAGATAAAGGTGAATAATGAGGAGAGTTGGTATGGAATAGATTG  
 TTACAAGGGAAGTGTATGTAACGCATACAAGGCTTGTATTTGGGAGCCGAG  
 TTTGGTGAAAAAAAACCTCAATTTACTCAGCTACTGAAGCAGCTTGCCTTGT  
 CTCTCAGTTGATGAAACTGTCAAAAACCAATCCAGACAACAGTTACAAAGCG  
 CACTACCACAACCCAAATAA

## Amino acid sequence of Tp4:

MSHLMNLPILVLKEGTDTSQGQAQIISNINACQAIVDCVKTTLGPRGMDKLIHTE  
 RDVTITNDGATVLKLLDITHPAASVLVDIAKSQDDEVGDGTTSVTVLAGELLNEA  
 KAFILDGISPVIIKYYREACQVALNLDKVAIHL SNKSSTDKKELLIKCAETTFNSK  
 LLSGYKTFFAKMVVEAVATLDEDLDEDMIGVKKVTGGSCEDSLLVKGVAFKKTF  
 SYAGAEQQPKKFVNPKILLNLELELKSEKENAEIVINNPQEYQKIIDAEYRIIFEK  
 LENA VKLGANVVL SKLP IGD LATQYFADKNVFCAGRVDENDLIRTSKATGASIQT  
 TLNNLSVDVLGTCGVFEEVQIGSERYNMFTDCKSAKTCTIVLRGGGQQFIDSE  
 RSLHDAIMIVRRATKCNILPGAGAIEMLLSTYLLHYS LNTINPTDSVNHVNCVNS  
 VNHVNGVTGVNKS LVGKRHIIMNGFAKALECIPRNLATNSGYNSNDLLSILRNKY

NQLEIVNGEIKVNNEESWYGIDCYKGSVCNAYKACIWEP SLVKKNSIYSATEAA  
CLVLSVDETVKNQSRQQLQSALPQPK

Figure 17: DNA and deduced protein sequence of Tp4. The complete ORF is 1740 nucleotides long with an ATG start and TAA stop codons. This ORF codes for a protein of 579 aa. BLAST searches revealed 100% identity with the T-complex protein of *Plasmodium* and *Babesia*. Analysis did not predict a signal peptide but showed the presence of a potential transmembrane domain.

Nucleotide sequence of Tp5 ORF:

ATGCCGAAAAATAAAGGCAAAGGAGGAAAGAACCGGAGACGCGGTAAAAAT  
GACAATGAAGGCGAAAAAAGAGAATTAGTCTTCAAATGGAAGATCAAGAAT  
ATGCTCAAGTTTTACGTATGCTCGGTAATGGCAGACTTGAAGCCTACTGTTT  
TGACGGCACTAAACGTCTTTGCCATATTAGAGGAAAGATGAGGAAGCGAGT  
TTGGGTAAATGCCGGCGATATTATTTTGGTATCGCTTAGAGATTTCCAGGAC  
AGCAAGGCTGACGTGATCGCAAAGTACACTGCTGAGGAGGCTCGTACTCTG  
AAGGCTTACGGCGAGTTGCCTGAAGCGACCAAAATCAACGAACTGACGTG  
TACGACGACGAGGCCGACAACTGCATTGACTTCCAGGACGTATCGTCTGAA  
TCAGAACCTGAGGATGAGTCACAAGAGGAGTCGGATTTTCGATATCGATGAT  
TTATAA

Amino acid sequence of Tp5:

MPKNKGKGGKNRRRGKNDNEGEKRELVFKMEDQEYAQVLRMLGNGRLEAYC  
FDGTRKRLCHIRGKMRKRVWVNAGDIILVSLRDFQDSKADVIAKYTAE EARTLKA  
YGELPEATKINETDVYDDEADNCIDFQDVSSSESEPEDESQEESDFDIDDL

Figure 18: DNA and deduced protein sequence of Tp5. The complete ORF is 468 nucleotides long with an ATG start and TAA stop codons. This ORF codes for a protein of 155 aa. BLAST searches revealed that Tp5 is the *T. parva* homologue of the eukaryotic translation initiation factor eIF-1A (BLASTX  $p(n)$ -value  $>10^{-65}$ ). Analysis did not predict a signal peptide or a transmembrane domain.

Nucleotide sequence of Tp7 ORF:

ATGACATCAAAGGACGAGACACCTGATCAGGAGGTCTACGCTTTTAATGCT  
GATATCTCCCAGCTTTTAAGCTTGATCATCAACGCATTTTATAGTAACAAGGA  
GATTTTCCTTCGTGAACCTCATTAGCAACGCTAGCGACGCACTGGAAAAAATT  
AGGTATGAGGCAATCAAGGATCCAAAGCAAATCGAGGATCAACCCGATTAC  
TATATCAGGCTGTATGCCGACAAGAACAACACCCTCACAATCGAAGATT  
CCGGTATTGGCATGACCAAAGCCGACCTCGTGAACAACCTCGGTACAATTG

CCAAATCCGGCACAAGAGCATTTCATGGAGGCACTGCAAGCAGGCTCGGAC  
 ATGTCAATGATCGGACAGTTTGGTGTGGTCTTCTACTCAGCATACCTGGTCTG  
 CAGATAAGGTGACAGTAGTGTCCAAGAACAACGCAGACGACCAGTACGTCT  
 GGGAGTCAACAGCCTCAGGCCACTTTACAGTGAAGAAGGACGACTCGCAC  
 GAGCCGCTCAAAAGAGGAAGTACTAGACTAATACTGCACTTGAAGGAGGACCAA  
 ACTGAGTACCTTGAGGAGAGAAGGCTGAAAGAGCTTGTTAAGAAGCACAGC  
 GAGTTCATTTTCATTCCCAATCTCGCTCTCAGTAGAGAAGACCCAGGAGACC  
 GAGGTCACTGACGACGAGGCAGAGCTAGACGAGGACAAGAAGCCCGAGGA  
 GGAGAAGCCCAAGGACGATAAGGTGGAGGACGTTACTGACGAGAAAGTGA  
 CCGACGTCACTGACGAGGAGGAGAAAAAGGAGGAAAAAGAAAAAGAAGAAG  
 AGGAAGGTCACCAACGTAACGCGTGAGTGGGAAATGCTTAACAAGCAGAAG  
 CCAATTTGGATGAGACTCCCGTCTGAAGTCACCAACGAAGAATATGCAGCG  
 TTCTACAAGAACTTAACCAACGATTGGGAAGACCACTTGGCCGTGAAACACT  
 TCAGCGTTGAGGGTCAGCTTGAGTTCAAAGCTCTACTGTTCTGTCCTCAAGAA  
 GAGCGCCGTTTGACATGTTGAGTCCCGCAAAAAGAAAAACAACATCAAGT  
 TGTACGTCAGACGCGTATTTATCATGGACGACTGTGAGGAGCTCATCCCGG  
 AGTGGCTTTCTTTGTGAAGGGTGTGGTAGACTCAGAGGACCTGCCCTTGA  
 ATATTTCTAGGGAACTCTCCAGCAGAACAAGATCCTCAAGGTCATCAGGAA  
 GAACTTGGTGAAAAAGTGCCTCGAGCTCTTCAATGAACTCACTGAGAAGAA  
 GGAGGACTTCAAGAAGTTCTACGAGCAGTTCAGCAAGAACCTGAAGCTGGG  
 AATCCACGAGGACAACGCTAATCGCTCAAAGATCGCCGAAGTGTGAGGTT  
 CGAGACAACCAAGAGCGGAGACGAACTCGTGTCACTCAAGGAGTACGTTGA  
 CAGGATGAAGAGTGACCAGAAGTATGTGTACTACATCACGGGAGAGTCGAA  
 GCAGAGCGTAGCCTCAAGTCCTTTCTTGAGACCCTGAGGGCTCGCGACTA  
 CGAAGTCCTGTACATGACTGACCCAATTGATGAGTACGCAGTTCAGCAGAT  
 CAAGGAGTTTGAAGGCAAGAACTCAAGTGCTGTACCAAGGAGGGCCTGGA  
 CCTTGATGAGGGCGAGGATGAAAAGAAGTCTTTGAAGCGCTCAAGGAAGA  
 AATGGAACCTCTTTGCAAGCACATCAAGGAAGTGCTCCACGACAAGGTGGA  
 AAAGGTCGTGTGTGGAACAAGGTTTACCGACTCTCCATGCGCACTTGTGAC  
 CAGCGAGTTCGGCTGGAGCGCGAACATGGAGCGTATCATGAAAGCACAAG  
 CTCTCAGAGACTCGTCCATAACAAGCTACATGCTGAGCAAGAAGATCATGG  
 AGATTAACCCGAGACATAGCATCATGAAGGAGCTCAAACTAGAGCTGCAA  
 ACGACAAAACAGATAAAACCGTCAAGGACCTAGTCTGGCTTCTCTACGACA  
 CAGCGCTCTTAACCTCAGGGTTTAACCTCGATGAGCCCACCCAGTTTGAA  
 ACAGGATCTACAGGATGATCAAGCTCGGACTCTCATTGGACGACGAGGAAC  
 ACGTAGAAGAGGACTCATCAATGCCGCCGCTGGATGAGCCCGTTGTGACT  
 CCAAAATGGAGGAAGTTGACTAA

Amino acid sequence of Tp7:

MTSKDETPDQEVYAFNADISQLLSLIINAFYSNKEIFLRELISNASDALEKIRYEAIK  
 DPKQIEDQPDYYIRLYADKNNNTLTIEDSGIGMTKADLVNNLGTIAKSGTRAFME  
 ALQAGSDMSMIGQFGVGFYSAYLVADKVTVVSKNNADDQYVWESTASGHFTV  
 KKDDSHPELKRGRRLILHLKEDQTEYLEERRLKELVKKHSEFISFPISLSVEKTQE  
 TEVTDDEAELEDKPKPEEEKPKDDKVEDVTDEKVTDVTDDEEEKKEEKKKKKRRK  
 VTNVTREWEMLNKQKPIWMRLPSEVTNEEYAAFYKNLTNDWEDHLAVKHFSV  
 EGQLEFKALLFVPRRAPFDMFESRKKKNNIKLYVRRVFIMDDCEELIPEWLSFVK



GVVDS EDLP LNISRET LQQNKILKVIRKNLVKKCLELFNELTEKKEDFKKFYEQFS  
 KNLKLG IHEDNANRSKIAELLRFETTKSGDEL VSLKEYVDRMKSDQKYVYYITGE  
 SKQSVASSPFLET LRARDYEVLYMTDPIDEYAVQQIKEFEGKKLKCCTKEGLDL  
 DEGEDEKKSFEALKEEMEPLCKHIKEVLHDKVEKVVCGRFTDSPCALVTSEFG  
 W SANMERIMKAQALRDSSITSYMLSKKIMEINPRHSIMKELKTRAANDKTDKTV  
 KDLVWLLYDTALLTSGFNLDEPTQFGNRIYRMIKLGLSLDDEEHVEEDSSMPPL  
 DEPVVDSKMEEVD

Figure 19: DNA and deduced protein sequence of Tp7. The complete ORF is 2166 nucleotides long with an ATG start and TAA stop codons. This ORF codes for a protein of 721 aa. BLAST searches revealed that Tp7 is the *T. parva* homologue of the heat shock protein 90 (hsp90) and showed 100% homology with a published partial sequence of *T. parva* hsp90 (accession # AAA30132). Analysis has predicted a potential signal peptide and a transmembrane domain.

Nucleotide sequence of Tp8 ORF:

ATGCTTGGAATCATGTCATGGGATCTAATCCCCCACATTAAAATTTTATC  
 ATCTGTTACATTCTTACATATTGCTAAAATGGAAGAAGTAGAAAACGTAAAAG  
 TCGACGCCTTGGAGCGTGTTGACACTGAGTCTGTCCTTAATTATGACACTGT  
 GTTAGAAAAGAAACCATTGCGCAGCAGTGTTGCCTCTTTCTTCAAAGATAC  
 AGTGCTGTTCTCGTAATATTAAGTCCGTGCTATTATTCACATTCACCTTTTGC  
 AGCAATAGCATTGTCATCAGGCAGAAGCGCAATCAGAAAGAACAGAGAACT  
 CCTGTCAGTCGAATTTGAAAAGCTTCAGTTCGATAATTCGTGACAATTAAG  
 GGAGAAAGGGAAGAGGACTTCCCCAAGATGGTAGCTGAAGTTCTTTACAAG  
 GTTGCAAGTCGAGTTTGACCCAAAAGAAGAGGCCTTGATCTACGTCCAGTTC  
 AATGACTTCAACAAGCAACACGACAAGAAGCACAACAATTACAGGCACAAG  
 AAGACCTCGTACACCAACTTCAGAAACAACCTTAATGATATAAACGAGCACA  
 ACGCAAAACCAACCTGTCGTACACCAAGAACATGAACCACTTCGGTGACA  
 TATCATCCAAGGATTTTATGAAGAGATACACCAAGAAAGTACTCTTGAACCTT  
 GCCAAAAGACCACGTGTCCACCTATAACAACAACAGACCAATGTCAGTTGAT  
 CTCAGAAGCCATGGTGTATTGACTCCAGTCAAGTGCCAAGAAGAAAATGAA  
 CTCTCATGGCCATACTCCGTAGTAGCAGTCGCCGAGTCATTCTGTTAAGAAG  
 ACATCACAAAAGACCGTATCCCTCAGCGAAAAACAATTAGTAGATTGCGTTA  
 CAGATAAGAAATCTGCAACAACCCATTCTTGGGTTACAAATACCTTAAGGA  
 CTTGGGTCTGTTTGAATCAGAACTCGTAGACAAATCCACAACCAAGTGCCC  
 AGCATTGGAAGGTGAAAGATTCAAAGTCCCATCATACTCATACTCATATGAG  
 CCAGATTTGGTGGCACTCTTGTTGAATGCAGGACCACTCACTGTACCAGTTG  
 CAGTGAGCGAGGATTGGCAATTCTACGCTGATGGAACCTTGGATGTATGCG  
 GTGCTGAATTGAACCACTTCTTGACCCTAGTAGGTGTCAGCTTTGACGAAAA  
 AGGCAATCACTGGATACTCAAAAACCTATTCCGGTGAAGGCTGGGGAAACAA  
 GGGATACCTACTGTTGACTCGCAATAGCAAGGAATACAAAGATGATTGTGG  
 ATTGACCTCCTTCGCAGTGACGCAGTTTAA

Amino acid sequence of Tp8:

MLGNHVMGSNSPHIKILSSVTFLHIAKMEEVENVKVDALERVDTESVLNYDTVLE  
 KKPLRSSVASFFKRYSAVLVILTAVLLFTFTFAAIALSSGRSAIRKNRELLSVEFEK  
 LQFDNFVTIKGEREEDFPKMVAEVLYKVAVEFDPKEEALIYVQFNDFNKQHDKK  
 HNNYRHKKTSYTNFRNNLNDINEHNAKPNLSYTKNMNHFGDISSKDFMKRYTK  
 KVLLNLPKDHVSTYNNNRPMMSVDLRSHGVLTPVKCQEENELSWPYSVVAVAES  
 FVKKTSQKTVSLSEKQLVDCVTDKKSANNPFLGYKYLKDLGLFESELVDKSTTK  
 CPALEGERFKVPSYSYSYEPDLVALLLNAGPLTVPVAVSEDWQFYADGTLDVC  
 GAELNHFLTIVGVSFDEKGNHWILKNSFGEGWGNKGYLLLTRNSKEYKDDCGL  
 TSFAVYAV

Figure 20: DNA and deduced protein sequence of Tp8. The complete ORF is 1323 nucleotides long with an ATG start and TAA stop codons. This ORF codes for a protein of 440 aa. BLAST searches revealed that Tp8 is a putative cysteine proteinase (BLASTX  $p(n)$ -value  $>10^{-30}$ ). Analysis has predicted a potential transmembrane domain but no signal peptide.

## 2.5 Protein Expression

### 2.5.3 Construction of Plasmids

The reading frames of *T. parva* candidate antigen genes were amplified by PCR using Taq polymerase (Promega, Madison, WI USA 53711) from the original full-length genes in their respective plasmids in pcDNA3. Both the forward and reverse primers contained restriction enzyme sites (Table 1). The PCR products were digested with the respective restriction enzymes, and ligated into bacterial His-tag expression vectors, pQE30 (Qiagen, 28159 Avenue Stanford, Valencia, CA91355) or PET28 (Novagen, Madison, WI 53719 USA). The plasmids were transformed into *E. coli* strain DH5 $\alpha$  (Life Technologies, Carlsbad, CA 92008, USA). All the plasmids were sequenced to ensure that they harbored no substitutions compared to the original genes. Purified plasmids were then used to transform competent BL21 DE3 bacterial cells.

Gene (Plasmid)	Forward Primer (Restriction Site)	Reverse Primer (Restriction Site)	Expressi on Vector	PCR Product (s) (kb)
Tp1 (pcDNA3)	GGATCCCCGGAAGAAAGAAG AGGAACTC ( <i>Bam</i> HI)	CTGCAGTTAAT TTTTGAGGTAAATTTT G ( <i>Pst</i> I)	pQE30	1.5
Tp4 (pcDNA3)	AATGTAGTTTTATCTAAATTGCC A ( <i>Bam</i> HI)	GAGGAGATAAG TTGAGAGCAACATC ( <i>Sal</i> I)	pQE30	0.4
Tp5 (pcDNA3)	GGATCCGAAATGGCGAAAAATA AAGGCAAAGGA ( <i>Bam</i> HI)	CTGCAGTTATAAATCA TCGATATCGAAATCT ( <i>Pst</i> I)	pQE30	0.6
Tp7 (pcDNA3)	GCCAAGAATTCGATGACATCAA AGGACGAG (gene internal <i>Bam</i> HI site)	GGCGCGGCCGCGTCA ACTTCCTCCATTTTG ( <i>Not</i> I, <i>Xho</i> I)	pET28b	2.0, 1.1

Table 5: Primers, restriction sites and vectors used for the cloning of *T. parva* antigens for protein production.

## 2.5.4 Expression of the proteins

Single BL21 DE3 bacterial colonies bearing the recombinant plasmids were isolated and cultured in 2XYT (formula) at 37°C to an OD<sub>600</sub> of 0.6, and protein expression induced by addition of IPTG to a final concentration of 1mM, and further cultured for 4 h. The cells were harvested by centrifugation at 4000 g for 20 min. Recombinant proteins were isolated by either the native or denaturing nickel-nitrilotriacetic (Ni-NTA) agarose according to the manufacturer's protocol (Qiagen, 28159 Avenue Stanford, Valencia, CA91355), dialyzed against PBS and stored at -20 °C. Purified proteins were checked on 12% SDS-PAGE gels and Western blot. Protein concentration was determined by the BCA Protein Assay reagent (PIERCE, Rockford, IL 61105, USA). Bacterial cells or purified proteins were applied to a 12% SDS-PAGE gel under denaturing conditions.

Proteins were electroblotted onto nitrocellulose sheets (Schleicher and Schull, Dassel, Germany). Mouse His-tag antibody (SIGMA) was used as the primary antibody, while anti mouse horse-raddish peroxidase conjugate (SIGMA) was used as the secondary antibody followed detection with 3,3'-Diamnobenzidine and hydrogen peroxide.

The reading frame of the segments Tp1, Tp4 and Tp5 amplified and cloned into the bacterial expression vectors harbored no substitutions compared to the original gene sequences. Recombinant protein containing His-tag were produced from all the constructs (Figure 21), and all can be determined by immunoblotting using His-tag antibody.

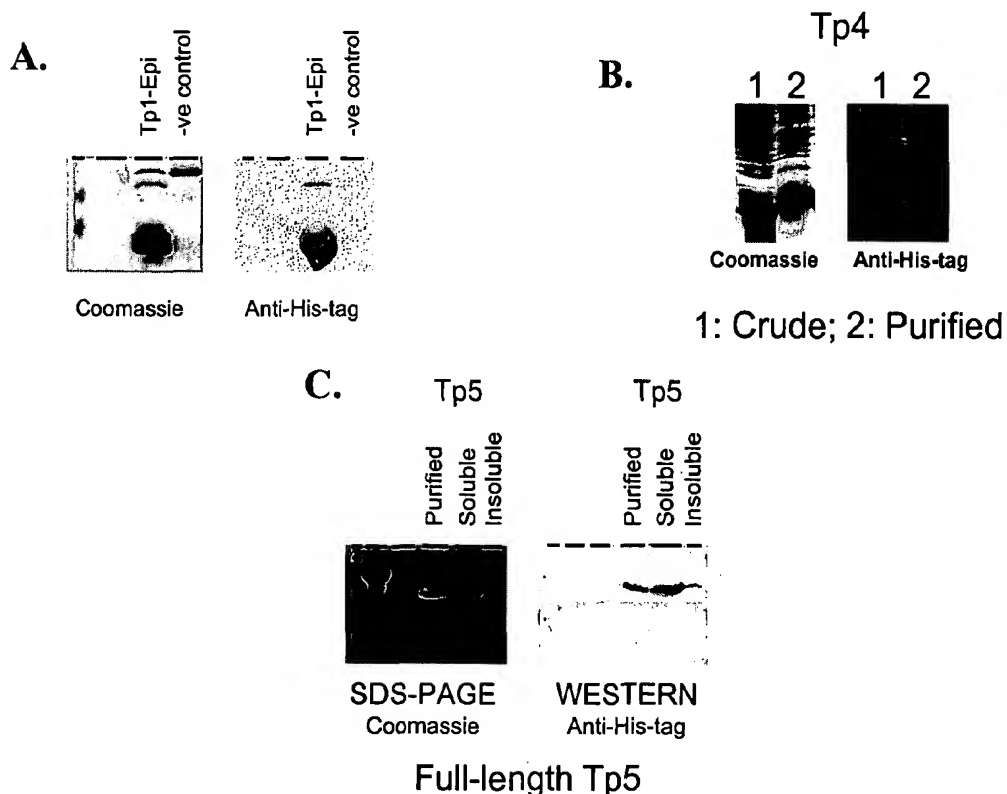
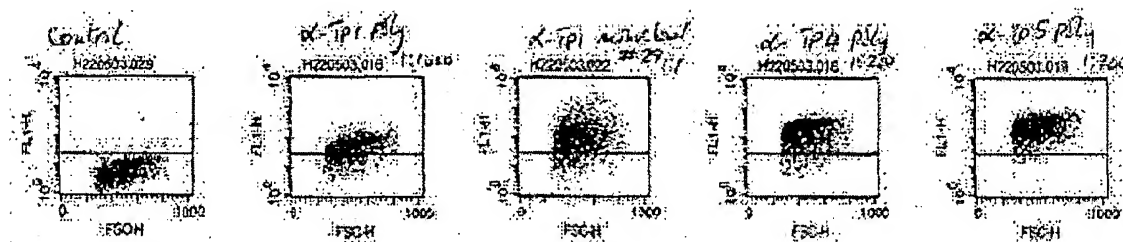


Figure 21: Expression of CTL target antigens, Tp1 (A), Tp4 (B) and Tp5 (C) proteins. Recombinant proteins were isolated by Ni-NTA agarose and ran on 12% SDS-PAGE gels followed by staining using coomassie blue or detection by Western blotting.

## Intracellular antibody staining



Antibody	% of <i>T. parva</i> infected cells stained
Anti-Tp1 polyclonal Ab	78
Anti-Tp1 monoclonal Ab	85.96
Anti-Tp4 polyclonal Ab	92.52
Anti-Tp5 polyclonal Ab	96.04
Control	0.32

Fig 22: Monoclonal and polyclonal antibodies raised against the Tp1, Tp4 and Tp5 proteins were used for staining of *T. parva* infected cells.

## Polymorphism in CTL target antigens

Specific Tp1 forward (IL #12588: 5'-ATG GCC ACT TCA ATT GCA TTT GCC-3' and reverse (IL #12589: 5'-TTA AAT GAA ATA TTT ATG AGC TTC-3') primers were designed and used in PCR to amplify a 430 bp region containing the BV115 CTL epitope from genomic DNA of *T. parva* (Kakuzi521, Kilifi KL2, Kilifi KL1, Kilifi BR305, Nyairo IL02, Nyairo IL17, D409 Tp Mariakani, Buffalo7344cl, Zambia 2 and Uganda) isolated from infected animals in different regions. PCR products were sequenced and their deduced amino acid sequences were compared with Tp1 (Figure 14). The result showed variations among the Tp1 studied. This strongly indicated that Tp1 is polymorphic.

Tp1	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIPAPPGVK
Kakuzi521	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIPAPPGVK
NyairoIL02	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIPAPPGVK
NyairoIL17	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIPAPPGVK
Kakuzi521	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIPAPPGVK
Buffalo7344cl	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEMLEMATKFNRLPKGVEIPAPPGVK
KilifiKL2	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEIIEMATKFNRLPKGVEIPAPPGVK
D409TpMariakani	MATSI AFAADPGFCYFLLI PAP--KPIFFKNDGDKFLRCVGYPKVKEEIIEMATKFNRLPKGVEIPAPPGVK
KilifiBR305	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEIIEMATKFNRLPKGVEIPAPPGVK
KilifiKL1	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEIIEMATKFNRLPKGVEIPAPPGVK
Zambia2	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEIIEMATKFNRLPKGVEIPAPPGVK
Uganda	MATSI AFAADPGFCYFLLI PGPD SKPIFFKNDGDKFLRCVGYPKVKEEIIEMATKFNRLPKGVEIPAPPGVK

Tp1	PEAPTPTPTTITPSVPPTIPTPTPSAPPTTPTTGLNFNLTQNKFMIGSQEVKLNITHEYEGVYEAHKYFI
-----	--

```

Kakuzi521      PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMIGSQEVKLNITHEYEGVYEAHKYFI
NyairoIL02     PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMIGSQEVNLTNITHEYEGVYEAHKYFI
NyairoIL17     PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMIGSQEVNLTNITHEYEGVYEAHKYFI
Kakuzi521      PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMIGSQEVKLNITHEYEGVYEAHKYFI
Buffalo7344c1  PEAPTPTPTITPSAPPTT-----PPTTPPKGLNFNLTQNKFMIGSQEVKLSITHEYDGVYEAHKYFI
KilifiKL2      PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMVGSQEVKLNITHEYDGVYEAHKYFI
D409TpMariakani PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMVGSQEVKLNITHEYDGVYEAHKYFI
KilifiBR305    PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMVGSQEVKLNITHEYEGVYEAHKYFI
KilifiKL1      PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMIGSPEVKLNITHEYEGVYEAHKYFI
Zambia2        PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMVGSQEVKLNIPHEYDGVYEAHKYFI
Uganda         PEAPTPTPTTITPSVPPTIPTPITPSAPPTTPPTGLNFNLTQNKFMVGSQEVKLNITHEYDGVYEAHKYFI

```

FIG. 23. Tp1 multiple sequence alignments. Amino acid sequence comparison of a portion (containing the HD6 CTL epitope) of Tp1 generated from *T. parva* isolates from different regions. Domains with variations are underlined.

# Identification of a vaccine candidate antigen recognised by *Theileria parva* specific bovin cytotoxic T lymphocytes

Simon P. Graham\*, Roger Pelle, Nyarhorwo J. Torukari, Mai Yamego, Rosemary M. Sanya, Elias Awino, John K. Nyarihi, Julius Ocaso, Anthony M. Muthland, Pierre van der Bruggen\* & Evans L.N. Taracha

International Livestock Research Institute, Nairobi, PO Box 30709, Kenya

\*Ludwig Institute for Cancer Research - Brussels Branch, Avenue Hippocrate 74, UCL 7459, B-1200 Brussels, Belgium

\*Correspondence - s.graham@ciglar.org

## ABSTRACT

Immunity against *Theileria parva*, the causative agent of East Coast fever (ECF), is mediated through lysis of schizont-infected cells by CD8<sup>+</sup> cytotoxic T lymphocytes (CTL). The identification of CTL target schizont antigens will aid the development of a sub-unit vaccine against ECF. A cDNA library was prepared from schizonts and used as a source of schizont antigens. Recognition of CD8-7 cells transiently co-transfected with pools of schizont cDNA and Bcl-2 class I/II cDNA by schizont specific CTL was assessed using an IFN- $\gamma$  ELISpot assay. A 543 amino acid secreted schizont antigen, termed Tp1, was identified and Tp1 transfectants shown to be lysed by CTL. Molecular and synthetic peptide approaches were used to define the CTL epitope on Tp1 as the 11mer, VGYPKVKEEM. CD8<sup>+</sup> T cell responses to this epitope observed following challenge of an immune bull with a lethal dose of *T. parva* sporozoites, suggested that Tp1 specific CD8<sup>+</sup> T cell responses might be protective.

## BACKGROUND/HYPOTHESIS

- East Coast fever, caused by the apicomplexan parasite *Theileria parva*, is an acute, usually fatal disease of cattle, which results in major economic losses throughout east and central Africa
- Immunity against *Theileria parva* has previously been shown to be mediated by MHC class I restricted CD8<sup>+</sup> cytotoxic T lymphocytes (CTL) directed against schizont-infected lymphocytes [1].
- It is hypothesized that identification of CTL target schizont antigens will aid the development of a sub-unit vaccine.
- Human melanoma specific CTL target antigens have previously been identified *in vitro* by measuring bioactive TNF- $\beta$  release from CTL co-cultured with COS-7 cells transiently transfected with melanoma cDNA and appropriate MHC class I cDNA [2].
- This strategy was adapted using an IFN- $\gamma$  ELISpot assay as a readout to detect *T. parva* specific CTL recognition of COS-7 cells transiently co-transfected with schizont and BoLA class I cDNA in an attempt to identify candidates for a CTL-based vaccine against ECF.

## METHODS & RESULTS

- Schizont specific CTL: A BoLA class I HD8 expressing Friesian bull (BV115) was immunised against *T. parva* (Muguga) by infection and treatment regime. Schizont specific, BoLA class I HD8 restricted CD8<sup>+</sup> CTL polyclonal lines and clones were generated using schizont infected cells as described previously [3].
- Screening schizont cDNA library for CTL target antigens: Purified *T. parva* (Muguga) schizonts used to generate unidirectional cDNA library in eukaryotic expression vector, pcDNA3. 1000 pools of schizont cDNA (50/pool) were co-transfected with HD8 cDNA into COS-7 cells. BV115 CTL fractions were co-cultured with transfectants and recognition assessed using an IFN- $\gamma$  ELISpot assay. Fig 1A shows the identification of 2 positive cDNA pools (B42 & B182) following screening of pools B1-B200. Positive cDNA pools were resolved to 48 pools of 10 cDNA using a 3-way matrix and screened with CTL to identify individual cDNA (Fig 1B). Confirmatory assays were finally performed with individual cDNA (Fig 1C).



•Identification of a CTL target antigen, Tp1: The single cDNA were sequenced and all positive cDNA encoded a 543 amino acid long hypothetical protein which was termed Tp1 (Fig. 2). The presence of a strongly hydrophobic domain was identified at the N-terminus by hydrophobic plot analysis suggesting it is a secreted protein. The gene encoding Tp1 was transcribed by the sporozoite, schizont and trophozoite life-cycle stages (Data not shown). BV115 CTL responded to 10 pools out of the 1000 pools prepared from the cDNA library. Tp1 was identified from 8 pools by resolution and single cDNA sequencing and from the remaining pools by Tp1 specific PCR.

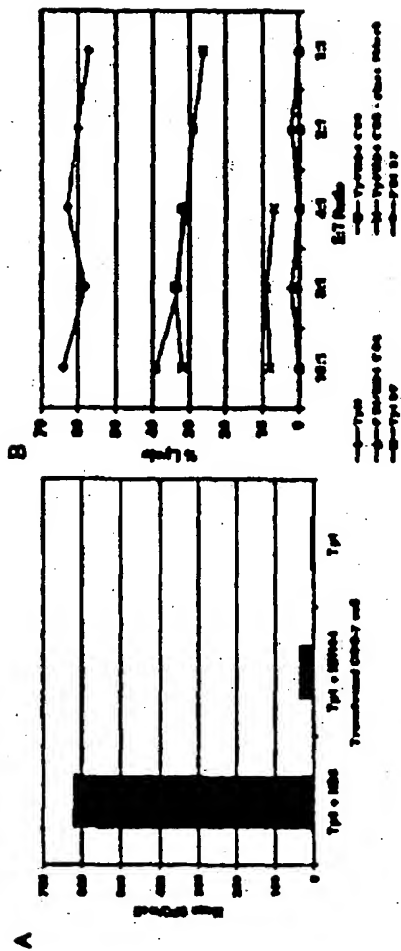
•Confirmation of Tp1 as a lytic target: HD6 restriction of Tp1 recognition was determined by assessing CTL recognition of cDNA co-transfected with HD6, an irrelevant BoLA cDNA (KON104) or alone (Fig. 3A). Using a <sup>51</sup>Chromium release assay, BV115 CTL were tested for their ability to kill COS-7 cells co-transfected with HD6 and Tp1 and also Tp1 transfected autologous skin fibroblasts (Fig. 3B). CTL lysed both HD6 expressing schizont-infected cells and Tp1 transfectants. Tp1 specific lysis was abrogated by the presence of a blocking monoclonal antibody against BoLA class I.

•Mapping of the HD6 restricted CTL epitope of Tp1: The region of Tp1 encoding the HD6 restricted epitope was mapped by Exonuclease III/Mung Bean nuclease digestion, followed by cloning of short overlapping fragments of Tp1 into pCDNA3 and screening recognition by BV115 CTL. This approach located the CTL epitope to a 60 amino acid stretch of Tp1 (Data not shown). An overlapping synthetic peptide library (12mer or overlapping by 2 amino acids) covering this region was used to pulse BV115 skin fibroblasts and screened with BV115 CTL clones (Fig. 4A). Responses against peptides 99, 10, 11 & 12 suggested that the epitope lay within the region FLRCVGYPKVKEEMLEMA. As possible 9, 10, and 11mer individual peptides spanning this region were screened by IFN- $\gamma$  ELISpot assay (Fig. 4B). This allowed the final definition of the HD6 restricted epitope as the 11mer VGYPKVKEEML.

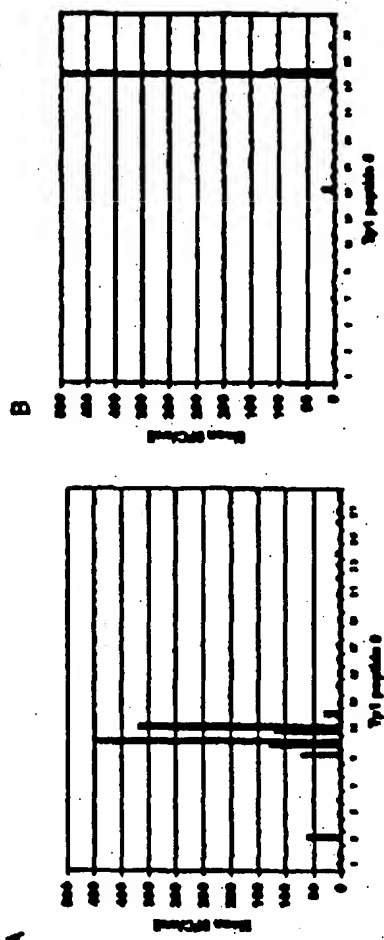
•Tp1 epitope polymorphism: The inability of BV115 CTL clones to recognise cells infected with the Marikebuni stock of *T. parva* or cells transfected with Marikebuni derived Tp1 (Fig. 5A) was investigated by comparing the deduced amino acid sequence of the original Marikebu stock Tp1 with Marikebuni Tp1. Whilst the two molecules were 95.9% identical, there was a double amino acid substitution at the C-terminus of the epitope (Fig. 5B).

[illegible]

Fig 2. Predicted amino acid sequence of Tp1. There is a putative signal sequence at the N terminus (underlined) with 3 putative cleavage sites (bold) with the most likely cleavage site between residue position 19 and 20 (pLA-G8) as predicted by von Heijne's method. The protein contains 6 potential N-glycosylation sites (red) suggesting a degree of post-translational modification.



**Fig 3. Confirmation of HDS restriction of Tp1 recognition and lysis.** CTL were cultured with COS-7 cells cotransfected with Tp1 and HDS cDNA, Tp1 and an irrelevant BclA class (KN104) or Tp1 alone, and recognition assessed using an IFN- $\gamma$  ELISpot assay (A). CTL lysis of isolated infected cells (Tp1) and Tp1 transduced COS-7 (co-transfection with HDS cDNA) and autologous skin fibroblasts (SF) was assessed using a  $^{51}$ Chromium release assay. To assess antigen specificity cells were transfected with an irrelevant antigen (PAM) and BclA class I restriction assessed by inclusion of a blocking monoclonal antibody against BclA class I (class I block) (B).



**Fig 4. Definition of the HDS restricted CTL epitope of Tp1.** Autologous skin fibroblasts (SF) pulsed with 28 overlapping peptides encompassing a 98 amino acid portion of Tp1 known to contain the CTL epitope were screened with two BV115 CTL clones (A). Twenty seven peptides comprising all possible 9, 10 & 11mers from the sequence FLNCVGYPKVKEEMLEMA were then used to pulse SF and screened with CTL clones (B). Recognition was assessed using an IFN- $\gamma$  ELISpot assay.

• Tp1 responses during protective immunity: Bull BV115 was challenged with a lethal dose of *T. parva* (Muguga) sporozoites and Tp1 specific CD8<sup>+</sup> T cell responses measured longitudinally by IFN- $\gamma$  ELISpot assay. The animal was solidly immune to challenge, developing no clinical symptoms. CD8<sup>+</sup> T cell responses were detected from day 8 post-challenge to the Tp1 epitope, not to a Tp1 11mer peptide that did not contain any known epitope and were sustained over the period of observation (Fig. 6). Significantly, the kinetics of the response was similar to that previously described for schizont specific CTL in efferent lymph and blood during a protective immune response [1].

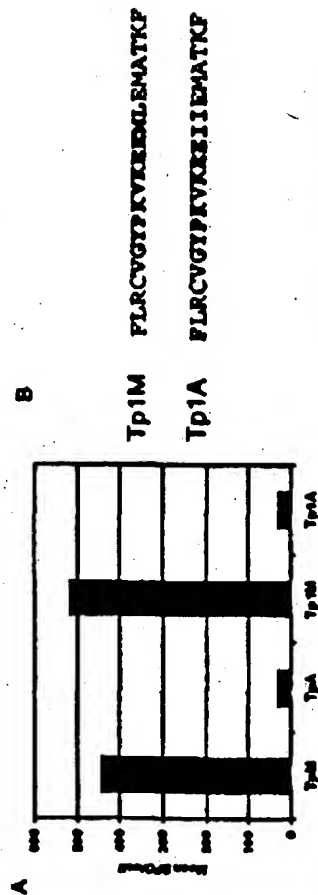


Fig 5. (A) Recognition of cells infected with Muguga (Tp1M) or Markubuni (Tp1A) sporozoites of *T. parva* and COS-7 cells transfected with HD8 and Tp1 derived from Muguga (Tp1M) or Markubuni (Tp1A) parasites by BV115 CTL was assessed by IFN- $\gamma$  ELISpot. (B) A comparison of the deduced amino acid sequence of the HD8 restricted epitope between Muguga and Markubuni Tp1 (bold) revealed a double amino acid substitution (red).

## CONCLUSIONS

- A sensitive high-throughput system has been adapted to facilitate the identification of bovine CTL target antigens. This has led to the first ever identification of a *T. parva* schizont antigen, Tp1, that is the target of CTL derived from an ECF immune animal.
- The CTL epitope has been defined and utilised to demonstrate increased numbers of Tp1 specific CD8<sup>+</sup> T cells during a protective immune response to a *T. parva* challenge.
- Tp1's vaccine potential should be evaluated in cattle under laboratory conditions using appropriate delivery systems and BoLA types.

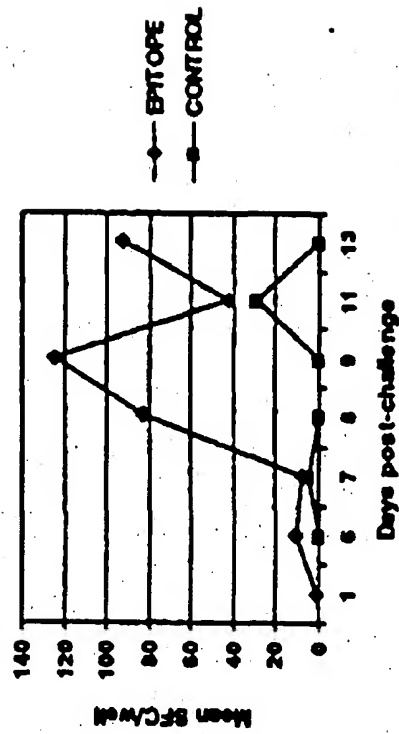


Fig. 6. Tp1 specific CD8<sup>+</sup> T cell responses of an immune bull following challenge with *T. parva* sporozoites. CD8<sup>+</sup> T cells were purified from peripheral blood at various time-points and incubated with the HD8 restricted Tp1 11mer synthetic epitope (epitope) or a Tp1 peptide which had not previously been shown to contain a CTL epitope (control) and response measured by IFN- $\gamma$  ELISpot assay.

#### REFERENCES

1. McGowan, D.J. et al (1994) Adoptive transfer of immunity to *Theileria parva* in the CD8<sup>+</sup> fraction of responding effluent lymph. *Proc. Natl. Acad. Sci. USA* 91, 1059-1063
2. De Plan, E. et al (1997) Identification of genes encoding for tumour antigens recognized by cytolytic T lymphocytes. *Methods* 12, 125-142
3. Goodwin, B.M. & Morrison, W.J. (1990) Techniques for the generation, cloning and characterization of bovine cytotoxic T cells specific for the protozoan *Theileria parva*. *J. Tissue Culture Methods* 11, 101-110

## **APPLICATION DATA SHEET**

### **Application Information**

**Application Number::**

**Filing Date::** July 14, 2003

**Application Type::** Provisional

**Subject Matter::** Utility

**Suggested Classification::**

**Suggested Group Art Unit::**

**CD-ROM or CD-R?::**

**Number of CD Disks::**

**Number of Copies of CDs::**

**Sequence Submission?::**

**Computer Readable Form  
(CFR)?::**

**Number of Copies of CFR::**

**Title::** EAST COAST FEVER VACCINE BASED ON CTL-  
SPECIFIC SCHIZONT ANTIGENS

**Attorney Docket Number::** 41860-189496

**Request for Early Publication?::**

**Request for Non-Publication?::**

**Suggested Drawing Figure::**

**Total Drawing Sheets::**

**Small Entity?::** Yes

**Latin Name::**

**Variety Denomination Name::**

**Patent Included?::**

**Patent Type::**

**Licensed US Govt. Agency::**

**Contract or Grant Numbers::**

**Secrecy Order in Parent Appl.::**

## **Applicant Information**

**Applicant Authority Typ ::** Inventor  
**Primary Citizenship::**  
**Country::**  
**Status::** Full Capacity  
**Given Name::** Evans  
**Middle Name::**  
**Family Name::** TARACHA  
**Name Suffix::**  
**City of Residence::**  
**State or Province of Residence::**  
**Country of Residence::**  
**Street of Mailing Address::**  
**City of Mailing Address::**  
**State or Province of Mailing Address::**  
**Country of Mailing Address::**  
**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor  
**Primary Citizenship::**  
**Country::**  
**Status::** Full Capacity  
**Given Name::** Jane  
**Middle Name::**  
**Family Name::** GLEW  
**Name Suffix::**  
**City of Residence ::** Nairobi  
**State or Province of Residence::**  
**Country of Residence::** KENYA  
**Street of Mailing Address::**

**City of Mailing Address::**

**State or Province of Mailing Address::**

**Country of Mailing Address::**

**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor

**Primary Citizenship::**

**Country::**

**Status::** Full Capacity

**Given Name::** Simon

**Middle Name::**

**Family Name::** GRAHAM

**Name Suffix::**

**City of Residence::** Nairobi

**State or Province of Residence::**

**Country of Residence::** KENYA

**Street of Mailing Address::**

**City of Mailing Address::**

**State or Province of Mailing Address::**

**Country of Mailing Address::**

**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor

**Primary Citizenship::**

**Country::**

**Status::** Full Capacity

**Given Name ::** Duncan

**Middle Name ::**

**Family Name ::** MWANGI



**Name Suffix::**

**City of Residence ::** Nairobi

**State or Province of Residence::**

**Country of Residence::** KENYA

**Street of Mailing Address::**

**City of Mailing Address::**

**State or Province of Mailing Address::**

**Country of Mailing Address::**

**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor

**Primary Citizenship::**

**Country::**

**Status::** Full Capacity

**Given Name::** Yoshikazu

**Middle Name::**

**Family Name::** HONDA

**Name Suffix::**

**City of Residence::** Nairobi

**State or Province of Residence::**

**Country of Residence::** KENYA

**Street of Mailing Address::**

**City of Mailing Address::**

**State or Province of Mailing Address::**

**Country of Mailing Address::**

**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor  
**Primary Citizenship::**  
**Country::**  
**Status::** Full Capacity  
**Given Name::** Roger  
**Middle Name::**  
**Family Name::** PELLE  
**Name Suffix::**  
**City of Residence::** Nairobi  
**State or Province of Residence::**  
**Country of Residence::** KENYA  
**Street of Mailing Address::**  
**City of Mailing Address::**  
**State or Province of Mailing Address::**  
**Country of Mailing Address::**  
**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor  
**Primary Citizenship::**  
**Country::**  
**Status::** Full Capacity  
**Given Name::** John  
**Middle Name::**  
**Family Name::** TONUKARI  
**Name Suffix::**  
**City of Residence::** Nairobi  
**State or Province of Residence::**

**Country of Residence ::** KENYA

**Street of Mailing Address::**

**City of Mailing Address::**

**State or Province of Mailing Address::**

**Country of Mailing Address::**

**Postal or Zip Code of Mailing Address::**

**Applicant Authority Type::** Inventor

**Primary Citizenship::**

**Country::**

**Status::** Full Capacity

**Given Name::** Matasuke

**Middle Name::**

**Family Name::** YAMAGE

**Name Suffix::**

**City of Residence::**

**State or Province of Residence::**

**Country of Residence::**

**Street of Mailing Address::**

**City of Mailing Address::**

**State or Province of Mailing Address::**

**Country of Mailing Address::**

**Postal or Zip Code of Mailing Address::**

## **Correspondence Information**

**Correspondence Customer Number::** 26694

**Phon Number::** (202) 962-4800  
**Fax Number::** (202) 962-8300  
**E-Mail Address::** magollin@venable.com

### **Representative Information**

**Representative Customer** 26694  
**Number::**

### **Domestic Priority Information**

<b>Application::</b>	<b>Continuity Type::</b>	<b>Parent Application::</b>	<b>Parent Filing Date::</b>
	<b>Continuation of</b>		
	<b>Continuation of</b>		
	<b>Continuation of</b>		
	<b>Continuation of</b>		

### Foreign Priority Information

Country::	Application Number::	Filing Date::	Priority Claimed::

### Assignee Information

**Assignee Name::** International Livestock Research Institute

**Street of Mailing Address::**

**City of Mailing Address::** Nairobi

**State or Province of Mailing Address::**

**Country of Mailing Address::** KENYA

**Postal or Zip Code of Mailing Address::**